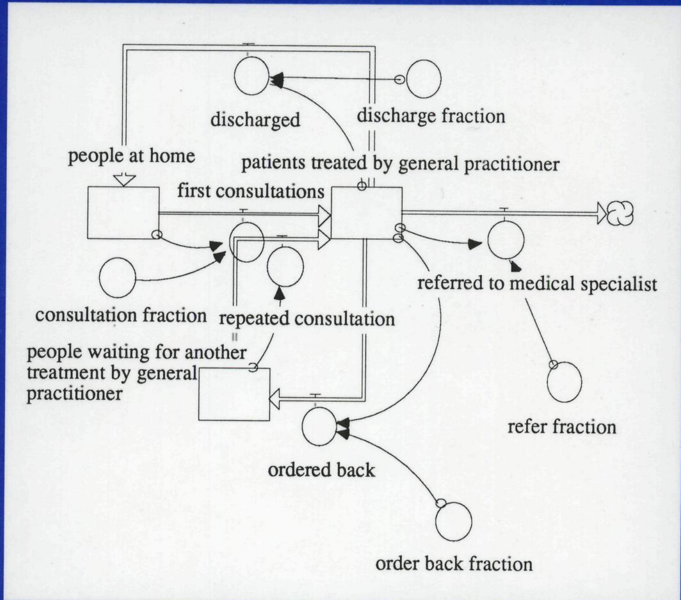
Participative Policy Modelling

# Participative Policy Modelling

## applied to the health care insurance industry




Luc Verburgh

Luc Verburgh

# PARTICIPATIVE POLICY MODELLING

## Applied To The Health Care Insurance Industry

## (PARTICIPATIEVE MODELBOUW
## Toegepast bij een zorgverzekeraar)

een wetenschappelijke proeve op het gebied van de Sociale Wetenschappen

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen in het
openbaar te verdedigen op
dinsdag 12 april 1994
des namiddags te 3.30 uur precies

door

Lucas Dagobertus Verburgh
geboren op 23 november 1961 te Leiden

Promotores: Prof. dr. J.H.G. Klabbers
Prof. dr. A.J.A. Felling

Co-promotor: Dr. J.A.M. Vennix

This study is part of the research program of the Nijmegen Institute for Cognition Research and Information Technology (Dept. of Mathematical Models and Decision Theory), from the University of Nijmegen.

Manuscriptcommissie:

Prof. dr. J. van Aken
Prof. dr. J. Geurts
Prof. dr. D. Post

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Verburgh, Lucas Dagobertus

Participative policy modelling. Applied at the health care insurance industry/ Lucas Dagobertus Verburgh. (S.I.:sn). Ill., fig., tab.

Proefschrift Nijmegen. - Met lit. opg. - met samenvatting in het Nederlands

ISBN: 90-9006953-4

Keywords: interactive simulation, strategic modelling, participative policy modelling

Druk: Benda, Nijmegen

# ACKNOWLEDGEMENTS

There are a number of people I wish to thank for their contribution to the process of writing my thesis. Without their support and friendship I might not have written a thesis.

I would like to begin thanking Jac Geurts and Jac Vennix for introducing me to the field of system dynamics modelling. Without their enthusiasm for the subject, I probably would not have pursued a job as research assistant in 1988.

It all started when Bert Felling was willing to have me interviewed for the job at a very short notice. Without his willingness to do so, I would not even have begun writing this thesis. In fact, someone else might have. I did not know Bert at the time that I started the research project. But I soon got to know him as the inspirational head of the University of Nijmegen's Department of Methodology, with firm values not only regarding what constitutes a good researcher, but also what constitutes a good person. His advice on research and personal matters has been a great help to me. Rather than saying I will miss his advice, I would like to say that I will miss him.

It was not until the second year in the process of doing the research, that I became acquainted with the ideas and values of another member of my thesis committee, Jan Klabbers. At first, I must admit, I was not always sure whether I did fully understand his theoretical notions and their implications on my research. I probably did not. It was only after lengthy dicussions that I became to appreciate his point of view. Looking back at those times and the impact those discussions have had on me and my thesis, I consider Jan to be a person who really helped me to develop as a scholar.

I am also very grateful that Jac Vennix was on my committee as well. Jac set the example for my research. The research he conducted a couple of years prior to the beginning of my research has been extremely useful. More importantly, Jac was a constructive person with whom to work - always critical, but in a positive way. Jac, I really enjoyed working with you, I wish we could have spent more together at the same university.

Doeke Post helped me dealing with the complexity of the health care system. Not only on a cognitive level by advising me on how the system works, but also on a social level by serving as an liason with the organization participating in program. Thank you Doeke.

Jan Gubbels and his broad methodological experience has provided me with methodological assistance and advice when I have needed it. Moreover, he adviced me not to expand the scope of my research, valuable advice for a doctoral candidate. I learned a great deal from Jan and enjoyed our conversations about research, money and cars.

Besides the members of my committee, many other people have contributed greatly to my graduate career. I want to thank the people from the Nijmegen Institute for Cognition Research and Information Technology (NICI) for having been a part of the department of Mathematical Models and Decision Theory and allowing me to do my research. I would like to thank Charles de Weert and Hil Weber in particular for making me feel a little bit a member of the NICI despite the fact that I was based at the Department of Methodology.

My colleagues at the Department of Methodology created an environment of support and friendship that allowed me to be myself and enabled me to complete succesfully the research program. Collegues whom I want to acknowledge in particular for the good times that I have had at the department and their contribution to my thesis are Jan Lammers, Piet Verschuren, Fred Wester, Ben Pelzer, Theo van de Weegen and Cecile Thijssen. Thanks.

Not yet acknowledged are the collegues who suported me in my annual preparation for the "Seven Hills Race". Ed Aarts, Chris Visscher, Marijke ter Voort, Piet Biemans and Mart van der Poel, thanks for your encouragement. Our weekly runs were a perfect way to replenish my energy reserves. Unfortunately, I never made it to the race, but I certainly enjoyed watching you run in it.

Some of my collegues of the Department of Methodology have become real friends. Vincent Peters, Hans van Kuppevelt, and Mart van der Poel, thank you for providing an ever available sounding board for my ideas, frustrations, anger, and need for having a chat. I will miss you.

Of the people who do not belong to University of Nijmegen faculty, but who have contributed greatly to both my professional and personal life, I would like to thank first my former collegues of the European Professional Educational (EPE) division of Arthur Andersen & Co. S.C. in Veldhoven. Not only because I had a great time in Veldhoven, but also because the EPE division was very helpful in my transfer to the Andersen Consulting office in The Haque. I also thank my collegues of Andersen Consulting who have shown interest in and concern for the status of my thesis and who have provided me with a very challenging and stimulating work-environment.

In addition, I wish to thank my friends for the good times that we have had over the last few years. Their 'distraction' made me realize that there is more to life than work alone.

Special thanks go to my best friend Robert Hooijberg, who has lived in the United States for more than 6 years now. Despite this separation, our friendship has proven to be very special. I also thank his wife Brenda Steinberg for being a friend.

My parents Carla and Bert Verburgh, and my parents-in-law, Erna and Roelof Gaspar, have supported both me and my family throughout my graduate career. I am very grateful.

Finally, I want to thank my wife Renée Verburgh-Gaspar and our children Maarten and Marleen, for their constant support and love.

# CONTENTS

# CHAPTER 1: INTRODUCTION

## 1.1 BACKGROUND OF THE STUDY

The way in which computer models have been applied to support the policy making process has changed over the last few decades. In the early days, computer models were mainly built by expert model-builders who eventually presented their model and recommendations in an impressive, difficult to read, research report. Because of the inaccessibility of the results and the difference in orientation and nature of expert modellers and policy makers (House, 1982), model-builders had difficulty in changing the policy makers' mind and thus hardly ever succeeded in having an impact on the policy making process (Brewer, 1973; de Man, 1987; Greenberger, Crenson, and Crissey, 1976; Geurts and Vennix, 1989; Meadows and Robinson, 1985; Watt, 1977).

To overcome these problems, it was recommended that relevant policy makers should be more involved in the modelling process (Carley, 1980; Greenberger et al., 1976; Geurts and Vennix, 1989; Meadows, Richardson, and Bruckmann, 1982; Vennix, 1990). By doing so, the model's accessibility would be increased for the policy makers themselves would be participating in the construction of the model, and thus partake in the process of building rather than being confronted with the results only. Because much of the understanding of the problem is generated in the process of model-building (Vennix, 1990), for instance in deciding what to include and what to exclude in the model, expert model-builders used to have a deeper understanding of the problem than the policy makers for whom the model was developed. Having the policy makers participate in the building process enables them to obtain as much understanding of the problem as expert builders usually have at the end of the model-building process. Another advantage would be that participation would lead to ownership and increased commitment, increasing the likelihood that the model will actually be used in the process of policy making. This alternative approach to model-building is called ***participative*** or ***interactive*** (as opposed to traditional) policy modelling (De Greene, 1982).

So far little systematic and objective empirical research has been carried out to examine whether participative policy modelling is able to live up to its expectations, that is, contribute to the policy making process. Most of the evaluations use self-evaluating questionnaires and/or outcome scores. Self-evaluating questionnaires provide some indication of whether participants have changed the way they look upon the policy problem. However, these questionnaires are ill-suited to examine where exactly in the participants' perspectives changes have taken place. Besides, self-evaluation is not the most objective way to assess whether people have changed. Outcome scores, which are often applied to assess how much students have understood of a model of a fictitious business, by contrast, are more objective in that the scores are not provided by the participants themselves but by the researcher using 'absolute standards'. Understanding and applying the knowledge that is modelled in the business-game correctly, results in higher scores and vice versa. So, the more the individual representation of the business problem resembles the model of the fictive business, the better the representation is, and the higher score the participant will get if (s)he acts accordingly. The reason why outcome scores cannot be used in the present study, is because we do not have an a priori model that can be regarded as a true representation of the reality our participants have to deal with; we do not have a scientifically validated model when dealing with complex, ill-structured problems. Consequently, we have no absolute standard that can help us to score the individual's model and their subsequent policy making behaviour. From a research point of view, this lack of standards obviously will have an impact on the present study's approach to assess the effects of participative policy modelling in that a different way of evaluating the effects must be sought for. However, one should bear in mind that because of this problem, because of the fact that a priori models containing all knowledge relevant for dealing with these particular kind of policy problem are lacking, the use of participative policy modelling is advocated. In other words, we will have to find a way around this problem in order to be able to assess the impact participative policy modelling is having on the policy making process.

The absence of a priori models for complex, ill-structured problems thus suggests that participative policy modelling can be of assistance by creating a model to represent and reduce this complexity. Complex, ill-structured problems are characterised by a large number of aspects and variables that are interconnected in various, often unknown, ways. Participative policy modelling can be used to collect and structure the knowledge that is required to represent and subsequently solve the problem. Moreover, not only do we have limited knowledge about these problems, but the knowledge that is available, is often scattered over many different sources (actors), who often are an expert on just one aspect of the problem and its solution (Mason and Mitroff, 1981). It is felt that participative policy modelling, because of its participative character, is able to integrate the different pieces of knowledge brought in by different experts or stakeholders, as a result of which a shared understanding of the complexity will be brought about. The evaluation of the extent to which and the way in which participation does indeed result in such a change in the participants' perception of a (policy) problem, is the present study's major objective.

One important consideration needs to be made at this point. Despite our choice for the application and evaluation of the participative policy modelling method in the context of complex, ill-structured policy problems, we do not mean to say that no other (participative) policy method exists that is capable of dealing with such complexity as well. However, rather than exploring which alternative approaches would, from a theoretical point of view, be applicable as well, it was decided to pursue in the Department's tradition of (participative) policy modelling and apply and evaluate the participative policy modelling method. The reason for this was not only because of the theoretical grounds for applying participative policy modelling in this particular context (cf. Geurts & Vennix, 1989; Klabbers, 1975; 1976; 1985; 1990; Meadows, 1980; Vennix, 1990; Vennix & Scheper, 1990), but also because of the considerable amount of experience that was available in applying system dynamics modelling to policy making (see for instance Geurts, 1981; Geurts et al, 1985; Geurts & Vennix, 1989; Klabbers, 1976, 1977; Vennix & Geurts, 1987; Vennix & Gubbels, 1988; Vennix, Gubbels, Post, 1986a, 1986b; Vennix, Gubbels, Post, and Poppen, 1988). However, in contrast to the amount of applications reporting the (successful) use of participative policy modelling in the context of policy making, hardly any of the studies report the use of evaluations to assess either the effects or the process by means of which these effects were brought about. Although some initiatives to evaluate the effects of participative policy modelling, computer-assisted learning environments, or computer-simulation, have been reported (Breuer & Kummer, 1990; Gould, 1989a, 1989b; Kim, 1989b), neither detailed descriptions of the approach taken nor presentations of the outcomes have been given so far with the exception of the Vennix (1990) study. The Vennix study, carried out as part of the 'Mathematical Models'-research program of the Nijmegen Institute of for Cognition and Information (NICI), used the so-called cognitive mapping technique to assess the changes in the way in which the participants looked upon the problem at hand. To arrive at the individual cognitive maps, Vennix had the participants write a policy document. Using the Axelrod (1976) coding procedure, individual cognitive maps were extracted from these policy documents (Vennix, 1990; Vennix and Scheper, 1990). The advantage of this approach lies in its increase in objectivity compared to the use of the self-evaluating questionnaires. Moreover, the cognitive mapping procedure provides a strong means to capture differences in both structure and content of the perception of the policy problem between and within those taking part in the participative policy modelling program.

Hence it was decided to carry out the present study in line with the Vennix (1990) approach. However, it was felt that at some point, the study's findings were to some extent indecisive, due to some practical limitations of the study. Consequently, some changes to the Vennix approach were made to overcome these limitations.

The first difference between the Vennix study and the present one concerns the people taking part in the program. In the Vennix study, the participants were all graduate students from various departments with only limited knowledge of the policy problem at hand. In the present study however, the participants work at a Regional Health Care Organization. They all have, some more than others, policy making experience and are knowledgeable regarding the policy problem at hand. It is this distinction that points to the underlying objective of the present study: evaluating the participative policy modelling technique in a real-world context, that is, applied at real policy makers rather than at members of a student population.

Secondly, the studies differ with respect to the kind of model that is being used. The Vennix study was based on an already existing econometric model which could not be changed by the participant. In the present study by contrast, a system dynamics model that can be altered by the participants themselves, is being used. As a consequence, the present study allows for more participation by the participants in the conceptualization of the model.

Thirdly, the criteria Vennix used to assess whether or not his participants had made any progress were derived from theories about the quality of policy documents (Hoogerwerf, 1984; Kraan-Jetten, 1986; Leeuw, 1983; Ringeling, 1985). Our study, by contrast, derives its criteria from the world of system dynamics (Forrester, 1961, 1969; Hanneman, 1988; Richardson & Pugh, 1981; Vennix, Smits, and Geurts, 1989). As a consequence, an assessment will be made of the degree to which people have changed their conceptualization in terms of system dynamics criteria; the degree to which they have been able to acquire a system dynamics perspective[1]. Although there is some overlap between the Vennix criteria and the criteria used in the present study, it is important to note that they stem from a different theoretical framework.

Having outlined the background of the present study in terms its main *subject* and *approach* in a general sense, we now need to focus on it in some detail.

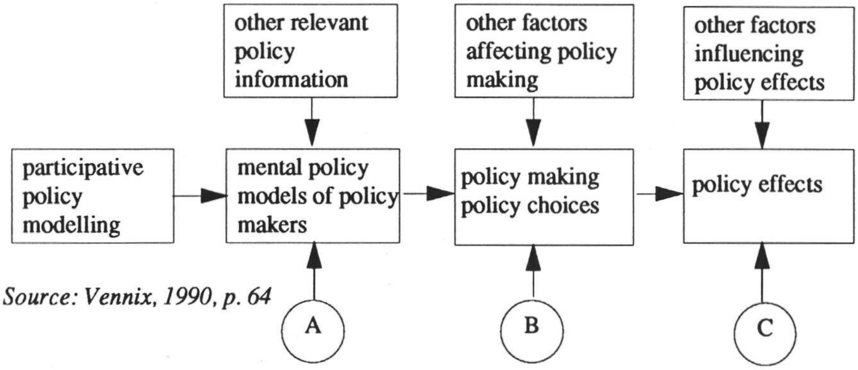
As far as the applicability of the participative policy modelling technique is concerned, it is important to note that we have in mind only one particular kind of policy problem: the so-called *complex, ill-structured* problems. The reason for this is not only because this lack of structure needs to be dealt with in order to be able to tackle the problem, but also because participative policy modelling seems to be capable of just doing that: providing a definition and structure of the complex, ill-structuredness that has to be dealt with by policy makers. In the next chapter, a more detailed description of the relationship between participative policy modelling and complex, ill-structured policy problems will be given. For now, it will suffice to say that the present study will be concerned with the evaluation of the effects of participative policy modelling *applied to complex, ill-structured problems.*

A second important consideration concerns the kind of effects we focus on in evaluating the impact of participative policy modelling. In line with Vennix (1990, p. 64), we distinguish among three potential points of measurement in evaluating the impact of participative policy modelling: the mental policy model of policy makers, the policy making (policy choices), and the policy effects.

---

1 What exactly these criteria are will be explained in Chapter Three.

Figure 1.1: Impact of computer policy models and points to measure their efects




*Source: Vennix, 1990, p. 64*

Figure 1.1 shows that we assume that policy modelling affects the policy making process because it changes the way people look upon the problem (their mental model). Further on, if the mental model changes (e.g. with respect to the problem definition) it is assumed that these policy makers will act differently (they will take different measures, make different choices), as a result of which different outcomes will be produced: the policy effects will be different. As Vennix (1990, p. 64-65) suggests, it would be nice to be able to examine the impact of participative policy modelling on the outcomes of the policy making process (i.e. measurements at point C). However, the existence of many potentially disturbing factors makes it difficult, if not impossible, to assess this empirically. The same argument holds for the policy making choices (i.e. measurement at point B) for there are many other factors affecting the policy making process along with the knowledge that is represented in the policy makers' individual mental map. Political considerations, for instance, play an important role in the process of selecting specific policy measures. Because of existence of many confounding factors and the limited (objective) empirical evaluation that has been carried out so far, we will concentrate on the conceptual rather than the instrumental impact of participative policy modelling, that is, the present study will be concerned with an evaluation based on measurements at point A. Note that even at point A we have to take into account other sources of information that may affect the participants' mental maps. As a consequence, the present study will focus on the question 'Does participation lead to changes in the participants' mental maps?' rather than 'Does p articipation change the policy makers' decisions?'. As such, we will primarily be concerned with the first stages of the policy making process, where problem definition and problem formulation/structuring (understanding the problem) are the major activities (Brewer and deLeon, 1983; Dunn, 1981; Mason and Mitroff, 1981; Quade, 1989). Not only are they the first stages in the policy making process, but '*in a sense formulation is the most important stage, for the effort spent restating the problem in different ways or redefining it, clarifies whether or not it is spurious or trivial and points the way towards a solution*' (Quade, 1989, p. 51). Or as Dunn (1981) puts it: '*...problem structuring is the most critical phase of policy analysis, since policy analysts often fail more often because they solve the wrong problem than because they get the wrong solution to the right problem*' (p. 98). The relationship between participative policy modelling and problem structuring will be discussed in more detail in the next chapter. For now it suffices to say that the first serves to support the second. Policy modelling can be of assistance to policy makers in their endeavours to deal with the problem by getting an idea of what the problem is, by restructuring their mental maps.

Having described the context within which the application of participative policy modelling is to be understood, attention can be focused on the major research questions of the present study. As stated before, the study's overall objective is *to assess if, where, and to what degree participation in participative policy modelling leads to changes in the participants' mental policy maps*. Two major questions can be derived from this:

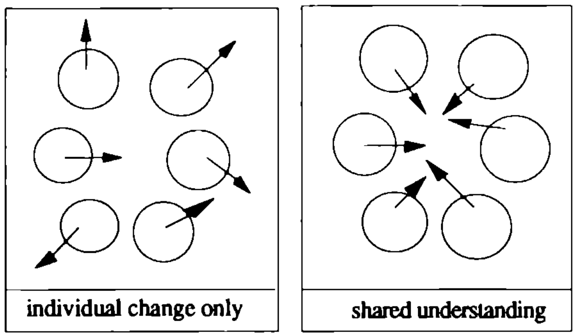
1. *Does participation change the individual mental maps (irrespective of the changes that take place in the mental maps of the other participants)?*

2. *Does participative policy modelling bring about a shared understanding among those taking part in it? I.e. is participative policy modelling able to increase the conceptual homogeneity among those participating in it?*

The first question focuses on how individuals change in the way they look upon the problem as a result of taking part in the participative policy modelling program. Using different variables in explaining the problem would be an indication of a change in an individual mental map.

However, learning and changing on an individual level does not automatically imply an increase in the participant's understanding of both the problem and the other participants. And since we are dealing with complex, ill-structured problems where knowledge about the problem is divided among many experts (Mason and Mitroff, 1981), communication and integration of this expertise is a major objective. Knowing about and understanding the problem on an individual basis is one thing, knowing about and understanding other people's perspective is something quite different. Obviously, inter-individual change cannot take place without individual change, but individual change alone does not guarantee an increase in homogeneity as depicted in figure 1.2.

Figure 1.2: Individual and inter-individual change




In discriminating between the individual and inter-individual effects of participative policy modelling, we come across another difference between the study conducted by Vennix and the present one. In the former, individual change was the major objective, whereas the present study will be concerned with individual and inter-individual changes in the way in which those who take part in the program look upon the problem at hand.

On the basis of the above mentioned distinction between individual and inter-individual changes, and in relationship to the support that participative policy modelling claims to provide in the policy making process, two hypotheses can be derived:

1. *Participative policy modelling changes the individual mental maps of those taking part in it.*

2. *Participative policy modelling increases the commonality (homogeneity) of the mental maps of those taking part in it.*

To assess the degree to which participative policy modelling is capable of living up to these expectations, criteria derived from the system dynamics perspective will be used.

However, to interpret these changes, a distinction between domain-specific and strategic change will be introduced. The reason for this is that in applying and adopting a system dynamics perspective not only changes in the knowledge base (the domain-specific content) may be brought about, but also changes in the way in which this content is organized and used in the process of dealing with complex, ill-structured policy problems. System dynamicists claim that the organization and use of the domain-specific knowledge has to be in line with the domain-independent system dynamics format in order to be applicable in the context of ill-structured policy problems. In other words, the changes in the conceptualization must be such that they, rather than showing a(n) (arbitrary) change in domain-specific knowledge only, illustrate the application of the system dynamics perception framework to the available domain-specific knowledge to arrive at a meaningful and appropriate (in light of the complexity confronted with) conceptualization of the problem. Application of the system dynamics format can be called 'strategic knowledge' because it guides the process of selecting and organizing the elements of the knowledge base (domain-specific knowledge) that need to be used to arrive at a useful representation of the problem.

In addition to examining whether the above two hypotheses hold or have to be rejected, the present study not only aims to critically review the theoretical framework used in this study on the basis of the empirical data gathered, but will also discuss (in an introductory mode, that is, as a starting point for future research) an alternative theoretical framework. This because it is believed that changing one's perspective may help to enrich one's understanding of the phenomenon under consideration. As will become clear after reading the concluding chapters of this study, such a change in perspective can be seen as a kind of 'closure' in that the change in perspective that is suggested obviously also affects (and has affected) our own position as researcher. The statements that we make do not only affect our object of research, but refer to ourselves as well.

## 1.2 PREVIEW OF THE STUDY

The present study focuses on the application of participative policy modelling to complex, ill-structured policy problems in order to support the policy making process. As such, three major components can be discerned: the policy making process, complex, ill-structured problems, and participative policy modelling. In Chapter 2, each of these three components will be described in more detail as well as the way in which they are related to each other. Moreover, a description will be given of the evaluations that have been carried out so far, to distinguish the present study from previous ones.

In Chapter 3, the major research questions of the present study will be examined in more detail, based on a description of system dynamics. This because participative policy modelling can be regarded as a participative version of system dynamics modelling. On the basis of the characteristics of system dynamics modelling, criteria will be developed along which the contribution participative policy modelling can make to the policy making process will be assessed. The result of this will be a theoretical or conceptual model of the present study, depicting the relationships between the dependent, independent, and potentially confounding variables.

In Chapter 4, a description will be given of the research design that was selected to assess the changes in conceptualization brought about at those taking part in the participative policy modelling sessions. The distinction between intended and actual design will be used to illustrate the impact the organization has had on the way in which the study eventually had been carried out. Prior to that, some of the methodological implications of notions of 'historicism' and 'situationality' introduced in Chapter 3, will be discussed. Following the description of the research design that will be used in the present study, an overview is given of the way in which the theoretical variables presented in Chapter 3 will be operationalized, on the basis of a pre- and posttest and a questionnaire. Finally, the time-table is presented to provide information on which activities have been carried out at what time, in the attempt to assess the effects of participative policy modelling.

The Dutch Health Care System will be focused upon in Chapter 5. It starts with an overview of the system in terms of competitiveness, costs, and health care insurance companies. This is followed by a description of the organization taking part in the participative policy modelling sessions. Next, the system dynamics model that was developed to speed up the construction of the model by the participants, will be presented. The model concerns the Dutch health care system from a Regional Health Care Insurance Organization's point of view. This because the client taking part in the participative policy modelling sessions is a Regional Health Care Insurance Organization.

The preliminary model (i.e. the model developed by us rather than the client), is divided into three subsystems: the processes surrounding general practitioners, the processes related to medical specialists, and the processes affecting the hospitals. The description will be followed by an overview of the model as a whole and a base-run, illustrating the dynamic behaviour of some of the variables of the model in the absence of any policy actions. Moreover, the effects of exogenous variables (such as ageing population and a policy measure) on the dynamic behaviour of the model will be presented to provide an understanding of both the preliminary model, and the activities carried out in the participative policy modelling sessions.

Chapter 6 is concerned with the construction of the variables. Firstly, the procedures which have been followed to transform the texts written by the participants into analyzable data will be presented. It will also discuss issues related to the validity and reliability of these procedures. Following the discussion of the cognitive mapping approach, the actual construction of the variables (i.e. the potentially specifying variables and dependent variables) will be presented.

The results of the current study are displayed in Chapter 7. A general description of the cognitive maps is presented first. This is followed by a description of the outcomes related to each of the present study's research questions. Next the values of the variables that are used as potentially confounding variables are presented to provide some information on how the program was perceived and the amount of time that was invested in it. On the basis of this description, an overview will be given of the most important effects the potentially confounding variables have on the relationship between participative policy modelling and conceptualization. Finally, comparisons with the Vennix study will be made to put the present study's findings into perspective.

Having described the outcomes, Chapter 8 will focus on their interpretation. It addresses the question whether or not the objectives have been met, how to account for the present study's outcomes, and what recommendations can be made regarding future studies. As part of this critical reflection, it will introduce and discuss an alternative theoretical framework that seems promising to account for the present study's findings. Note that this alternative framework should be regarded as a first step towards a better understanding of the relationship between participative policy modelling and policy making (support) rather than the introduction of a full-fledge alternative theoretical model.

Following the critical review of the study's outcomes and theoretical framework, a brief summary of the main conclusions will be given in Chapter 9.

# CHAPTER 2: POLICY MAKING AND ILL-STRUCTURED PROBLEMS: PARTICIPATIVE POLICY MODELLING

## 2.1 INTRODUCTION

To understand why the present study aims at evaluating the impact of participative policy modelling on the participants' mental maps while claiming to discuss its impact on the policy making process, the relationship among policy making, ill-structured policy problems, and participative policy modelling needs to be elaborated upon. To do so, a description will be given of the policy making process and its distinctive phases. Relating these phases to the characteristics of ill-structured policy problems, a list of features, that techniques contributing to the policy making process must have, will be provided. This is followed by a description of the participative policy modelling programme itself, to arrive at an understanding of how the characteristics of participative policy modelling match the requirements that stem from the description of ill-structured problems. Finally, in order to portray the characteristics of the present approach to evaluate the effects of participative policy modelling, comparisons with other related studies will be made.

## 2.2 POLICY MAKING VERSUS POLICY ANALYSIS: PROCESSES AND PHASES

Prior to talking about techniques that contribute to the policy making process, one should have a clear understanding of what is meant by policy making. The reason for this is that to appreciate the importance of the process of problem structuring (the process that will be focused upon in the present study), one should understand its context, that is, its relationship to the other aspects of the policy making process. Moreover, to avoid potential confusion regarding the terms to be used to refer to this context, account should be taken of the ways in which the terms 'policy making' and 'policy analysis' are currently being used. Eventually, a decision has to be made whether to use the term 'policy making' or 'policy analysis' to refer to the process of which problem structuring is an element.

Regarding the term 'policy making', three definitions can be discerned. These definitions differ in the way in which the elements of planning and action are thought to be related to each other (Sleegers, 1991). The first definition concentrates on planning as opposed to action (e.g. Kuypers, 1986, p. 79). According to the second definition, policy making is primarily concerned with 'doing something about the problem' (Dunn, 1981, p. 46; Dye, 1978, p. 3), whereas the third definition attempts to include both planning and acting (Rosenthal, Van Schendelen & Ringeling, 1986, p. 191). Brewer and deLeon (1983) call this combination of planning and acting the 'policy process' rather than making a distinction between policy making (acting, deciding) and policy analysis (planning, preparing).

For each of these three definitions of policy making, a different relationship with policy analysis exists:

If policy making is defined as primarily a planning activity (that what precedes a particular policy), policy analysis can no longer be used to refer to the process of producing knowledge *in* the policy processes. In other words, the purpose of policy analysis then is not to help policy makers to resolve the issues they face (because that is how policy making is defined), but to analyze already implemented policies (Kuypers, 1980, p. 358). The relationship between policy making and policy analysis thus can be expressed as follows: policy making is the process by means of which policy makers plan for action (for policy) and policy analysis is the analysis carried out to study the products of policy making.

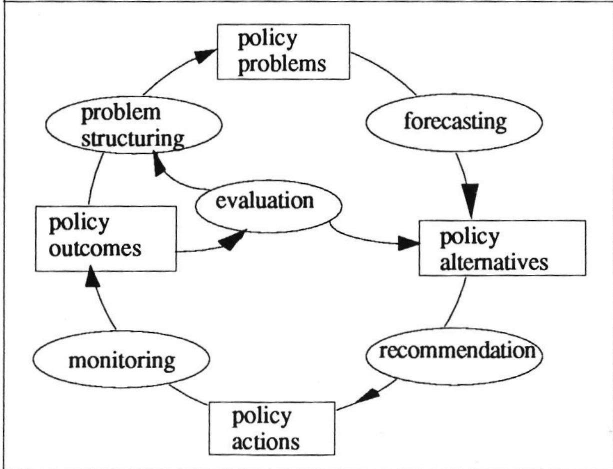
If policy making is defined as primarily 'doing something about the problem', policy analysis can be defined as the process of producing knowledge *in* the policy process. As a consequence: *"Policy analysis is essentially a cognitive process, while policy making is a political one. Many factors other than methodology shape the ways that policy analysis is utilized by policy makers: the structure of political power; the political feasibility of recommended alternatives; time and resource constraints; the form and content of information; and the characteristics of policy makers themselves"* (Dunn, 1981, p. 46). Policy analysis thus refers to the cognitive process that helps policy makers in the process of policy making, that is, to help them to *"acquire a deeper understanding of sociotechnical issues and to bring about better solutions"* (Quade, 1989, p. 4). It (policy analysis) is the planning that precedes action (policy making).

If policy making includes both planning and doing something about the problem, then policy analysis can either be regarded as the analysis of the process of policy making (meta-analysis), or the analysis of current policy (Kuypers, 1980).

Another option would be to abandon the terms policy making and policy analysis altogether and employ the term policy (or policy process) as suggested by Brewer & deLeon (1983) and Geurts & Vennix (1989).

For our purposes, it is not important to determine which of the definitions is most valid, as long as the reader has a clear understanding of what the *activities* are for which participative policy modelling claims to be of any help. And although the discussion whether these activities should be called policy, policy making or policy analysis is a very interesting one, it need not be resolved to appreciate the present study. However, to avoid confusion with respect to the terms that are used in the present study, we suggest that the third definition of policy making, where policy making includes both the planning and the actual actions, should be adopted. The reason for this is that it has the advantage that one term suffices to refer to the whole process. Consequently, the present study will employ the terms 'policy making' or 'policy process' to refer to both the planning (cognitive) and action-related activities of the policy process.

Figure 2.1: The policy making model (based on Dunn, 1981, p.48)



Considering the process of policy making in more detail reveals that the process can be divided into stages that are linked together and may have to be repeated several times (Brewer & deLeon, 1983; Dunn, 1981; Geurts & Vennix, 1989; Quade, 1989). Note that *"although to explain the process of analysis* [in our terminology: policy making] *it is necessary to take a sequential approach, in practice the order in which the topics I am calling stages are considered is relatively unimportant."* (Quade, 1989, p. 50). Or as Dunn (1981, p. 50) puts it *"Although the central direction of policy analysis conforms to the series of clockwise transformations we have described, the actual process of analysis may at times involve forward and backward movements between informational components"*. The model used to explain the policy making process is an analytical model which does not necessarily need to be similar to the daily practice of policy making. However, the model is useful in that it assists in depicting the components that play an important role in the process of policy making.

To understand *where* exactly participative policy modelling aims to contribute to the policy making process, we need to have a closer look at the policy making process. For this we will use the model developed by Dunn (1981), as depicted in figure 2.1.

It is shown in the figure that the model consists of four stages representing the four major *policy-informational components*. These components serve to make visible the outcomes of the so called '*transformation processes*'. They represent the four major issues that are dealt with in the policy making process:

1. *What is the policy problem?*
2. *What alternatives do we have? What can we do about it?*
3. *What actions are we going to take? What are we going to do about it?*
4. *What are the outcomes? Is it effective?*

To answer these four questions, a policy maker has to carry out transformation processes. In other words, (s)he has to go from one component to the other in order to arrive at the last question. By doing so, the questions are not answered in isolation, but in relationship to each other. The policy alternatives that are being generated to move from the first to the second component (and question) are related to the answers given to the first question to make sure that the alternatives are relevant, that is, related to the problem as defined in the problem definition component.

The distinction between transformation process and policy-informational components provides a first (although partial) answer to the question where participative policy modelling aims to contribute to the process of policy making. It is believed that participative policy modelling will be of assistance in the (transformation) process of going from one component to another. In other words, it claims to help the policy maker to use the information encompassed in one component to move towards an answer to the next question. The results of this contribution materialize in the quality of the answer, that is, the quality of one of the four components: policy outcomes, problem definition, problem alternatives, and policy actions.

### *The policy making model in more detail*

As said before, the actual policy making process may or may not conform to the logical reconstruction (the model) of the policy making process. Policy makers may start anywhere in the policy making process rather than beginning at the logical point of departure which is *"sensing that there is a situation or set of external conditions which, once experienced, gives rise to dissatisfaction, uneasiness, and a "felt recognition" that something is wrong"* (Dunn, 1981, p. 135). Or as Brewer and deLeon (1983, p. 36) point out: *"conflict or tension must exist between familiar patterns of behaviour and expectation and one's environment"*.

However, because we believe that a clear description of the policy making process is best served by following the logical order within the policy making process, the description of the process will start with the transformation process that transforms the feeling that something is wrong (based on the outcomes of previous policies or autonomous developments) into a clear description of the problem. The activities of this transformation process are called **problem structuring** or problem formulation (Brewer & deLeon, 1983, p. 42). In particular when dealing with complex, ill-structured problems, beginning the policy making process by structuring the problem is very important: *"An analysis must begin with problem formulation A major pitfall is the failure to allocate the total time intelligent, so that a sufficient share of it will be spent in deciding what the problem really is* "(Brewer & deLeon, 1983, p. 42).

Not only is problem structuring the first phase of the policy making model, it also is one of the most important ones: *"Problem structuring, which is that phase in the process of inquiry where analysts grope toward possible definitions of a problematic situation, is no doubt the most crucial but least understood aspect of policy analysis* [in our terminology: policy making]" (Dunn, 1981, p. 98). Or as Mason & Mitroff (1981, p. 29) put it: *"First, for problems of complexity, . the problem-defining phase is critical It may, in fact, be the most critical phase of all"*.

The importance of problem structuring is first of all based on its ability to help policy makers to avoid committing errors of the third kind.[2] It helps them to solve the right problem. Developing a number of policy alternatives and consequently selecting and implementing one of them is useless unless the policy problem (the result of the problem structuring process) is accurately representing the original feeling of what was wrong, the problematic situation, as Dunn (1981, p 109) calls it If there is no congruity between the problematic situation and the problem as defined by the problem structuring activities, there is a chance of solving the wrong problem As Quade (1989, p. 24) puts it: *"A problem well-put is a problem half solved "*

Secondly, problem structuring is important because of the impact it has on the other stages of the policy making process (Gusfield, 1981; Kalff, 1989). Based on the definition and/or formulation of the problem, alternative solutions are being developed and evaluated (ex ante evaluation).[3]

The third reason why problem structuring is an important aspect of the policy making process is, as will be explained in more detail in the section concerning the characteristics of ill-structured problems, because many of the problems policy makers face are complex and ill-structured. To deal with this complexity, policy makers can and often do reduce it at an early stage. For instance, by defining the problem in terms of well-known and well-structured problems or by describing the problem in terms of their solutions. By doing so, they make themselves guilty of what has been called 'premature closure'; in their attempt to simplify the complexity, alternative formulations are being excluded and alternative solutions are being denied. If, by contrast, problem structuring is taken seriously, it is much more likely that policy makers will formulate the problem without jumping at already existing descriptions and definitions.

Finally, problem structuring is important because of the fact that there are large differences in the way in which policy makers conceptualize the problem. *"Each expert has his own perspective of the problem and hence a different problem representation..."*

---

2 "One of the most popular paradigms in mathematics describes the case in which a researcher has either to accept or reject a so called null hypothesis In a first course in statistics the student learns that he must constantly balance between making an error of the first kind (that is, rejecting the null hypothesis when it is true) and an error of the second kind (that is, accepting the null hypothesis when it is false). practitioners all too often make errors of a third kind· solving the wrong problem" (Raiffa, 1968, p 264)

3 "The generation of policy alternatives builds directly from the preceding steps of the initiation stage [problem structuring The individual alternatives are a function of all that has gone before " (Brewer & deLeon, 1983, p 64)

"The most important criteria for deciding on a solution strategy is the problem representation." (Premkumar, 1989, p 560)

(Premkumar, 1989, p. 569). To arrive at a generally agreed upon plan of how to approach the problem, problem structuring is essential. Without having the individual participants restructure their original conception of the problem, it is impossible to overcome the differences in conceptualization and integrate the individual conceptions into one generally agreed upon problem representation.

What exactly the *activities* are that belong to the process of problem structuring will be described in detail in the section called 'problem structuring'. For now it suffices to say that problem structuring is the process which restructures the original conception of the problem (the problematic situation) into a well-structured and specified description. Or as Quade (1989, p. 51) puts it:

> "...to clarify the objectives, to discover the major factors that are operative, and to get some feel for the relationships among them. At the start these relationships may be extremely hypothetical, for empirical information and knowledge are likely to be in short supply; but the attempt at clarification will help to make the logical structure of the analysis appear."

Having structured the problem and thus having arrived at a specific description of the policy problem, policy makers need to start thinking about alternatives to solve the problem. In order to arrive at the second component of the policy making process, policy alternatives have to be formulated and evaluated. In other words, one has to think about the potential effects of the alternatives in terms of alleviating (or maybe even solving) the problem. The transformation process which covers these activities, is called '**forecasting**'. Forecasting is important because *"it is only by understanding the future that we are able to control it"*. And *"it* [forecasting] *helps policy makers foresee and avoid unanticipated negative consequences of public policy ..."* (Dunn, 1981, p. 140).

To explore the potential effects of particular alternatives, quantitative techniques such as time-series analysis, regression analysis and quantitative models can be used. However, particularly when dealing with complex, ill-structured problems, quantitative techniques are of limited usefulness. Human judgement and intuition need to be used when a unambiguous theoretical or empirical model does not exist. Techniques that cán be used in situations like these, are for instance scenario writing (Quade, 1989), theory mapping, causal modelling (Dunn, 1981) and participative policy modelling (Geurts & Vennix, 1989). For a more detailed description of possible techniques, the reader is referred to Brewer & deLeon (1983), where an overview of the techniques that can be employed, is presented in a section called 'A survey of general procedures'.

In the next stage of the policy model, alternatives are being compared to arrive at policy **recommendations**. Techniques that are often employed for this are cost-benefit analysis (in this approach the total monetary costs and total monetary benefits are being quantified)[4], and cost-effectiveness analysis (comparing the alternatives by quantifying their total costs and effects; in contrast to cost-benefit analysis, which tries to measure all relevant factors in one common unit of value, cost-effectiveness analysis uses two different units of values (Dunn, 1981, p. 250)). Following this evaluation, a decision has to be made whether or not to implement one (or more) of the policy alternatives. The part of the process where the actual implementation is taking place, is called **policy action**.

Once the policy alternative has been implemented, policy makers are not done. It is important that they **monitor** the outcomes to see whether the objectives are being met. As such, monitoring provides information about the relationship between the policy-program operations and their outcomes. To monitor the program, techniques such as social systems accounting, social experimentation, social auditing and social research cumulation are being used. Since, as will be explained in more detail later, the major

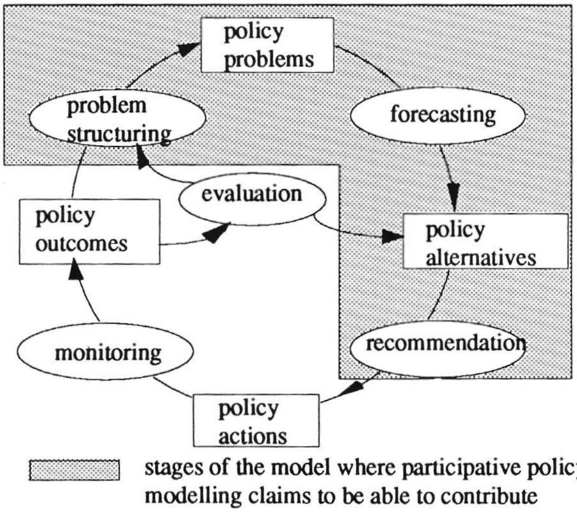
---

4 For a comprehensive description of cost-benefit analysis see, for example, Mishan, E.J. (1976). Cost-Benefit Analysis. New York, F.A. Praeger.

contribution of participative policy modelling to the policy making process will certainly not be at the monitoring stage, a more explicit description of monitoring as part of the policy making process is not required (for a more detailed description, the reader is referred to Dunn (1981, Chapter 8)). If the policy outcome is being perceived as a failure, policy makers may decide to start the whole process all over again, as a result of which they may again begin with the problem structuring stage.

Having discussed the stages of the policy model, a second answer to the question of which aspects of the policy making process participative policy modelling aims to be of any help can be provided. Now that we can distinguish among the various stages of the policy making process, a distinction can be made between the early stages, that is, the stages of problem structuring, forecasting, and recommendation, and the later stages, where the actual decision and implementation is taking place. Participative policy modelling claims to be most helpful to the early stages, the stages that precede the actual decision and implementation of one or more of the policy alternatives.

Figure 2.2: The policy making model (based on Dunn, 1981)

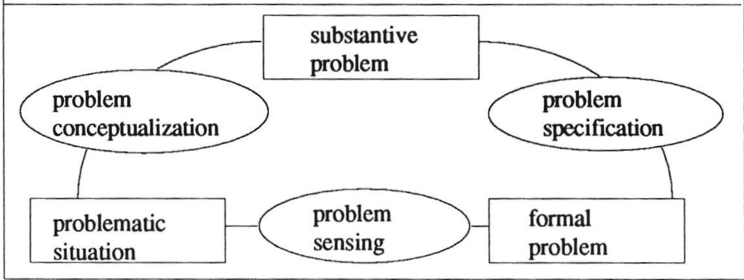



The reason why participative policy modelling claims to be of any help to these developmental stages in particular is twofold. Firstly, the very own characteristics of the technique itself make it more appropriate for contributing to the early stages of the policy making process. Why exactly will be explained in more detail in the section describing the participative policy modelling program in more detail. Secondly, the developmental stages are more appropriate because we restrict ourselves to so-called complex, ill-structured policy problems. Their lack of structure need to be dealt with first, before policy makers can start thinking about the selection and implementation of a policy alternative. How and why the features of ill-structured problems result in the decision to support the 'early' stages of the policy making model will be explained in more detail in the section on complex, ill-structured problems.

Having said that participative policy modelling aims to contribute primarily to the problem structuring stage of the policy making process, a more detailed description of this stage should be given. Examination of the problem structuring stage will reveal that the stage is made up of three sub-stages (problem sensing, problem conceptualization, and problem specification) and will help to explain why assumptions play an important role in the way in which the problem is being perceived and defined, in particularly when dealing with ill-structured problems. By doing so it will give an idea of why problem structuring and ill-structured problems are closely related to each other.

### *Problem structuring in more detail*

Figure 2.3: problem structuring (based on Dunn, 1981, p.107)




Distinguishing among interdependent phases within the problem structuring process is not typical for the model we use. Mason & Mitroff (1981), for instance, also use a threefold distinction, although their terminology differs somewhat from our model ('problem sensing', 'problem defining', and 'formal modelling').

The problem structuring process starts with ***problem sensing***, the recognition of a discrepancy between the situation as it is perceived (the 'is') and the 'desired' situation (the 'ought'). However, problem sensing (being aware of such a discrepancy) alone is not enough for a discrepancy between 'is' and 'ought' to become a problematic situation, that is, to become part of the policy making process. Two additional aspects are required. First of all, the belief that something can be done about the discrepancy. If this assumption of 'attainability' is violated, the felt discrepancy will remain a problem, but will not become a *policy* problem (Nelissen, Geurts, de Wit, 1986, p. 23). Secondly, it is important that the discrepancy is being recognized by the organization. It has to appear on its agenda. *"Failing to gain the organization's attention, an individual's recognition of the problem will not be shared by the organization nor will it be acted upon."* (Brewer & deLeon, 1983, p. 38). So it is only when both conditions are being met, that a discrepancy is perceived as a problematic situation.

In moving from a problematic situation to a substantive problem, conventional (native) language is being used to describe the problem in its most basic and general terms. The process by means of which this transformation is brought about is called ***'problem conceptualization'***. It concerns the decision of which conceptual framework to use to describe and to define the problem; the decision which frame of reference is most appropriate to interpret the aspects of the problematic situation in a meaningful way. Examples of possible frameworks are economics, sociology, and political science. If the problem is viewed as an economic one (the resulting substantive problem then is conceptualized as an economic problem), factors such as production and distribution and market prices become important elements. However, if the problematic situation is conceptualized as, say, a sociological one, it may be approached in terms of power, interest groups, elites, and social strata. In other words, <u>the resulting substantive problem is a general description of the problem in terms of one particular conceptual model.</u> Mason & Mitroff (1981) call this process 'problem defining', where different macro point of view (that is, conceptual frameworks) are being compared in the attempt to ensure that the problem is defined correctly.

The last stage of the problem structuring model consists of a more detailed description and definition of the problem (***problem specification***) in terms of the conceptual framework previously selected. Its purpose is to arrive at a formal description of the problem, which, as Mason & Mitroff (1981, p. 24) put it, *"consists of operationalizing* *and examining in detail a single micro view of the problem using the results from previous problem defining phase that appear to be "most fruitful," "promising", or "relevant" for the particular problematic situation at hand "*

Note that the second stage of the problem structuring process in particular, is affecting the *kind* of explanations and causes that are brought about in the problem specification. The selection of, say, an economics perspective, automatically results in economic causes for the policy problem at hand. The impact the conceptual framework has on the causes and solutions that appear in the policy making process is one of the reasons why, as will be extrapolated upon later, a multi-disciplinary (methodological) approach is preferred to a mono-disciplinary one.

To illustrate the active part that policy makers (or analysts) play in defining the nature of the problem itself (in particular when dealing with ill-structured problems), and to show how the choice of a "worldview" is affecting their formulation of the problem, look at the following, classic example:

> "The manager of a large office building received an increasing number of complaints about the elevator service in the building. He engaged a group of engineers to study the situation and to make recommendations for improvements if they were necessary. The engineers found that the tenants were indeed receiving poor service and considered three possible ways of decreasing the average waiting time. They considered adding elevators, replacing the existing ones by faster ones, and assigning elevators to specific floors. The latter turned out to be inadequate and the first two were prohibitively expensive to the manager. He called together his staff to consider the report by the engineers. Among those present was his personnel director, a psychologist. This young man was struck by the fact that people became impatient with a wait which seemed so short to him. On reflection he became convinced that their annoyance was due to the fact that they had to stand inactive in a crowded lobby for this period. This suggested a solution to him which he offered to the manager, and because it was so inexpensive the manager decided to try it. Complaints stopped immediately. The psychologist had suggested installing large mirrors in the walls of the lobbies where people waited for the elevators." (Ackoff, 1969, pp. 431-432).

The example shows that if the problem is conceptualized as an "engineering" problem, it becomes a problem of elevators, whereas, if it is conceptualized as a "behavioural" problem, people are at the core of the problem. As a consequence, conceptualizing the problem as an engineering one automatically results in solutions that focus on the elevator (speeding up the elevator, putting in more elevators and so on). Defining it as a people's problem by contrast, leads to solutions that seek to modify the people involved in the problem rather than the elevator. Note, that the example should not be interpreted as showing that one approach is better than the other, but that *"problems are cognitions and recognitions - products of our conceptual imagination. As a result, different analysts often can and do conceptualize problems in different ways."* (Mason & Mitroff, 1981, p. 26). One and the same problematic situation may result in two or even more substantive problems, depending on the conceptual framework that is used to define it. Problem structuring thus can be viewed as a process of human interpretation, and it is because of this interpretative nature that assumptions (and assumed knowledge) can play an important role. Differences in background, experiences, and training allow people to differ in the way in which they perceive one and the same problematic situation. Because they use different conceptual 'devices' to arrive at description of the problem that makes sense to them, it is likely that they differ in the eventual description and definition of the problem at hand. Particularly when the policy making process is carried out by more than one individual (say, within the context of organizations), differences in perspective and thus description of the situation or problem may become an important issue:

> "...in the early 1960s, President John F. Kennedy dispatched two senior officials, a Foreign Service officer and an Army general, to Vietnam to provide him with fresh and realistic estimates of the situation there. The general's glowing assessment, when matched against the civilian's gloomy pessimism, prompted Kennedy to remark: "Were you two gentleman in the same country?"[5]. Up to some limit (clearly surpassed in the example), diversity of viewpoint may be a positive thing: responsible decision making requires consideration of numerous plausible action before selecting only those having promise. In highly complex situations, exceeding the limit can lead to trouble.[6]" (Brewer & deLeon, 1983, p. 89)

Consequently, one of the most important aspects of the policy making process is to get the participants, who are often experts on one particular aspect of the problem and who often have different (educational and experiential) backgrounds, to work together and communicate their expertise.

> "When dealing with a multi-faceted problem with the aid of a variety of experts of different backgrounds, perhaps the most important requirement in the interest of an efficient use of these experts is to provide an effective means of communication among them. Since each of the participating experts is likely to have his own specialized terminology, a conceptual alignment and a real agreement as to the identity of the problem may not be easy to achieve, and it becomes almost imperative to construct a common frame of reference in order to promote a unified collaborative effort." (Helmer, 1966, pp. 17-18)

As already presented in the introductory chapter of the present study, participative policy modelling claims to affect the individual participants' mental map, that is, their individual representation of the problem, *and* aims at changing the way in which the participants' mental maps relate to each other. More precisely, participative policy modelling aims to bring about an enrichment in the individual mental maps of the participant and to establish a common frame of reference, that is, an overlap between the individual mental maps: a shared understanding, a common frame of reference (Quade, 1989, p. 197).

Having considered that multiple interpretations of reality are very common in the problem structuring phase, account needs to be taken of the reasons why participants differ so much in their perception of the problem. Certainly, differences in personality traits, past experiences, and training play an important role, but not all the differences can be explained on the basis of differences among the 'interpreting actors' of the problematic situation. The problematic situation itself, that is, the policy problem prior to transformation into a substantive problem, allows for more or less variety in interpretation. For instance, if the policy problem is well-structured, it is not likely that multiple interpretations will appear. On the other hand, if the policy problem at hand is complex and ill-structured, different perspectives are bound to emerge. Since participative policy modelling claims to contribute to the policy making process when applied to ***ill-structured problems*** (problems that lack a definite expression (Kalff, 1989), and are likely to be interpreted in more than one way), a more detailed description of ill-structured problems will be given. Following the description of the features of ill-structured problems, an overview will be given of the requirements that policy methods for ill-structured problems must meet to qualify as useful. Based on these requirements, an assessment will be made whether participative policy modelling is well-suited to tackling ill-structured problems.

---

5 As reported in M. Halperin, Bureaucratic Politics and Foreign Policy, Washington D.C.: The Brookings Institution, 1974, p. 171.

6 For more on this issue, see for instance Allen, (1978, p. 21); Brewer & deLeon, (1983, p. 98) and Carley (1980, p.13)

## 2.3 Ill-STRUCTURED PROBLEMS

Most problems can be classified according to their degree of structure. Depending on the amount of information that is available about the discrepancy between the 'is' and the 'ought', problems can be classified as being well-structured, semi-structured, or ill-structured (VanGundy, 1988). Well-structured problems are problems where all the information required to close the gap between the 'is' and the 'ought' is available. Ill-structured problems, by contrast, are those where the policy maker is provided with little or no information on the best way to develop a solution. And semi-structured problems, finally, fall in between the other two types.

To describe the differences among the ill-structured, moderately structured, and well-structured problems in more detail, consider how they vary on the elements they have in common (Dunn, 1981; Geurts and Vennix, 1989):

Figure 2.4: Three classes of policy problems (Dunn, 1981, p. 103)

| | structure of problem | | |
|---|---|---|---|
| elements | well-structured | semi-structured | ill-structured |
| decision makers | one or few | one or few | many |
| alternatives | limited | limited | unlimited |
| utilities (values) | consensus | consensus | conflict |
| outcomes | certainty or risk | uncertainty | unknown |
| probabilities | calculable | incalculable | incalculable |

The figure shows that there are at least five aspects that play an important role in the assessment of the degree to which policy problems are structured. These common elements are: the number of policy makers that are involved, the number of policy alternatives that seem relevant, the degree to which there is consensus on the values that are at stake, certainty of the policy outcomes, and the degree to which the probabilities can be calculated.

Applying these five elements to the category of ill-structured problems results in a description of the ill-structured problems according to which they *"typically involve many different <u>decision makers</u> whose <u>values</u> are either unknown or impossible to rank in a consistent fashion. Their <u>policy alternatives</u> and <u>outcomes</u> may be unknown such that <u>estimates of risk and uncertainty</u> are not possible."* (Dunn, 1981, pp. 104-105; underlined by the author). As a consequence: *"The problem of choice is not to uncover known deterministic relations, or to calculate the risk or uncertainty attached to policy alternatives, but rather to define the nature of the problem."* (Dunn, 1981, p. 105).

Based on figure 2.4, Geurts & Vennix (1989) conclude that ill-structured problems are both <u>cognitively</u> and <u>socially</u> complex. Ill-structured problems have a cognitive complexity because of the large number of variables that are involved and the limited amount of knowledge that is available about these variables and their relationships. As far as the relationships between the variables are concerned, features such as feedback loops, delays and non-linearity make it extremely difficult to understand let alone predict the behaviour of those problems. Others have recognized the cognitive complexity inherent to ill-structured problems as well:

> "Overlapping interactions among numerous elements, positive and negative feedback control loops, nonlinear relationships, and continuous structural changes in social systems. These characteristics largely account for the astonishing diversity of social systems and behavior. Our limited intellectual apparatus, however, prompts us to seek simply-ordered regularity. Our images are poor proxies for reality. Analyses frequently reflect these defective images, and so, too, do our policies." (Brewer & deLeon, 1983, pp. 88-89)

Regarding the social complexity of ill-structured problems, Geurts & Vennix (1989) point out that because ill-structured problems usually have more than one person or organization partake in the policy making process, it is likely that conflicting values, competing claims and different perspectives on the problem arise:

> "With increased complexity come increases in the number and diversity of system interpretations, in part because of the biased and distorted views affected individuals bring with them to the problem context." (Brewer & deLeon, 1983, p. 89)

Having said that ill-structured problems have both a cognitive and a social complexity, we can return to the question of why policy makers differ so much in the way in which they see the problem when dealing with ill-structured problems, to provide a more detailed answer.

Ill-structured problems almost inevitably call for different interpretation of the (problematic) situation because of the above mentioned cognitive and a social complexity: *"Analysts and decision makers have different motivations, different life experiences, divergent purposes, and often different intellectual capabilities and training, all of which suggests reasons identical factual situations so often produce extremely different alternatives for action"* (Brewer & deLeon, 1983, p. 89). Moreover, differences in perception also arise because of the impossibility to take into account all the factors that are related to the problem. One has to reduce the compiexity by simplifying it within the context of ones background and values Or as stated by Mason & Mitroff (1981, p. 18):

> "Complex problems depend on assumptions because it is not humanly possible to know everything of importance about a problem of complexity prior to the taking of action."

However, in reducing this complexity, people differ, due to differences in life style, purpose, experience etc.

So what we are faced with when dealing with ill-structured, complex problems, are situations where there is no uniform framework that can be used to interpret it (Klabbers, 1990), no existing theory that can be used to organize and structure the felt disagreement about the situation, because the situation under consideration is unique and history provides no guidance (Kalff, 1989). There are only competing frameworks (Rein, 1976) and the multiple views and values of the actors that are involved, have only little in common. As a consequence, one of the first steps to be taken when confronted with ill-structured problems, is to *'design a new process involving a larger than usual number of managers and experts with different organizational loyalties who can commit themselves to irreversible decisions <u>on the basis of a rich and joint perception of a problem and its context'</u>* (Kalff, 1989, p. 55; underlined by the author). Prior to discussing what this 'new process' should look like, that is, how the above mentioned features of ill-structured problems affect the way in which they should be approached, one question needs to be answered; the question of whether ill-structured problems are indeed a kind of problem that policy makers regularly are being confronted with. In other words, is the present study concerned with a kind of policy problem that is relevant to policy makers and the practice of policy making?

This question can be answered affirmatively; many of the policy problems policy makers face are indeed ill-structured[7] in nature:

> "Today few of the problems policy makers face are truly problems of simplicity or of disorganized complexity." (Mason & Mitroff, 1981, p. 9)

Not only that, because as Dunn (1981, p. 105) puts it, these ill-structured problems are also very critical in nature:

> "One of the main tasks of policy analysis is therefore the resolution of ill-structured problems. The reasons why ill-structured problems are so critical for public policy analysis have been aptly summarized by a number of social scientists." (Dunn, 1981, p. 105)

## 2.4 HOW TO APPROACH ILL-STRUCTURED PROBLEMS

Before criteria for the design of policy making methods for ill-structured problems can be developed, a summary of the area where policy making and ill-structured problems converge in the process of dealing with ill-structured problems need to be given. The reason for this is that our criteria need to be related to the features of both policy making and ill-structured problems.

To summarize what has been said about the <u>policy making process</u>, recall that the description followed a model in which a distinction was made between components and transformation processes. The major components were policy outcomes, policy problems, policy alternatives, and policy actions. In order to move from one component to another, one subsequently had to carry out the transformation processes of problem structuring, forecasting, recommendation, and monitoring.

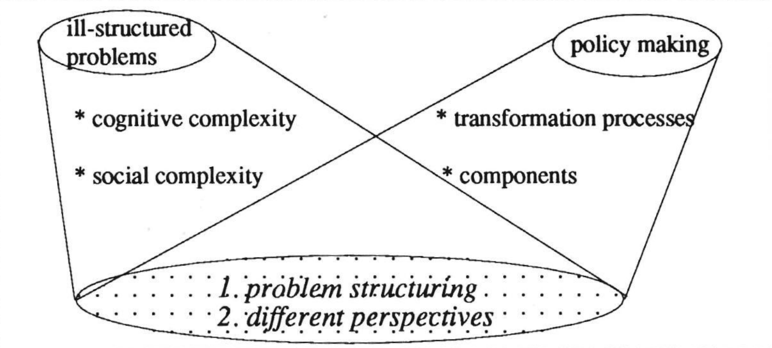
The question of where exactly participative policy modelling aims to contribute to the policy making process was answered in two different ways. Firstly, it was pointed out that in general, participative policy modelling aims to offer help at the transformation stages. Secondly, it was shown that participative policy modelling aims to contribute to the problem structuring process in particular because of the characteristics of the method itself (as will be explained in the next section) and the features of ill-structured problems. Recall, people tend to disagree on the definition of the problem as a result of the social and cognitive complexity inherent to ill-structured problems: *"In short, an ill-structured problem is one for which it is <u>not</u> relatively easy to obtain consensus concerning the basic definition of the problem."* (Mason & Mitroff, 1981, p. 30). Participative policy modelling claims to help the individuals involved in the process of defining the problem to establish a common perspective of the problem on the basis of which the rest of the policy making process can be carried out. This points to one of the reasons why problem structuring was said to be of importance: because of its impact on the stages that follow the structuring of the problem; errors being made in the problem structuring process for instance, are carried on throughout the rest of the policy making process.

As mentioned before, <u>ill-structured problems</u> have both a social and cognitive complexity. The cognitive complexity is partly due to the large number of variables that seem to be related to the problem and the relationships among those variables. The social complexity stems from the number of people (stakeholders, experts) that are involved in the process of dealing with the problem. Because of the social and cognitive complexity, participants are likely to differ in the way in which they perceive (and thus represent) the problem. Relating the ill-structured problem's feature of complexity to the policy making process's feature of stages in general and the stage of problem structuring in particular, it

[7] Different terms used to refer to the same kind of problems are 'wicked' problems (Webber and Rittle, 1973), 'messes' (Ackoff, 1974), and 'squishy' problems (Strauch, 1975).

should become clear that the differences in perspectives concern the problem structuring stage. In other words, participative policy modelling as a policy making method should take account of the way in which the participants differ regarding the way in which they describe and define the problem at hand (and help dealing with these aspects). In dealing with ill-structured problems, problem structuring is one of the most important stages of the policy making process, and it is very much likely that the people that take part in structuring the problem, differ in their original conception of the problem. The way in which this is reflected in the requirements that policy methods for ill-structured problems must meet, will be discussed in the next section. The relationships between the policy making process and ill-structured problems can be summarized as follows:

Figure 2.5: Ill-structured problems and policy making




***Policy making methods dealing with ill-structured problems***

When dealing with ill-structured problems, policy making is primarily concerned with the initial phase or phases of the policy making process. These initial phases have been referred to as articulation (Mitchell, 1980), problem definition (Klabbers, 1983), problem structuring (Dunn, 1981), or problem setting (Schön, 1983). The main activity that is carried out during this first stage of policy making in the context of ill-structured problems, is to shape order, create knowledge, and improve the mutual awareness among the participants (Klabbers, 1990). Based on the description of the activities and expected outcomes of the problem structuring stage, and the discussion of the context within which it is carried out, a number of implications for policy making methods can be derived. Policy methods dealing with ill-structured problems should be[8]:

**1. Participative**

One of the first implications for the design of policy making methods is that there must be a broader participation of affected parties (Kalff, 1989; Premkumar, 1989). Since the knowledge required to solve the problem is likely to be dispersed among a variety of experts and since there is a diversity of parties at stake, broad participation is required to ensure that all the relevant information to solve the problem is drawn upon:

> "Since the relevant knowledge necessary to solve a complex problem and also the relevant sources necessary to implement the solution are distributed among many individuals, the methods must incorporate the active involvement of groups of people." (Mason & Mitroff, 1981, p. 16)

8 Most of the guidelines for policy making methods are based on Mason & Mitroff, 1981.

**2. Integrative**

Because, as described above, people tend to differ in their interpretations of ill-structured problems. They are likely to disagree on what factors are crucial and on how they are related to each other. The knowledge and assumptions they have regarding the problem at hand tend to differ; their mental maps contain different elements due to differences in cognitive background and value systems. However, in spite of the individual differences at the beginning of the policy making process, that is, in the perception of the problem, a commonly agreed upon set of assumptions (or knowledge) about the problem at hand should be arrived at to deal effectively with the consequences of ill-structured problems. If those involved in the policy making process do not agree on the basic formulation of the problem, they are bound to disagree on the steps that need to follow.

> "A unified set of assumptions and a coherent plan of action are needed to guide effectively policy planning and strategy making... Something else is needed to bring this diverse but relevant knowledge together in the form of a total picture." (Mason & Mitroff, 1981, p. 16)

The major objective of the problem structuring stage thus not only is to enrich the individual participant's understanding of the problem, but also to arrive at a commonly agreed upon formulation and definition, that is, to bring about an integrated perspective, a 'shared understanding of the problem at hand' (De Geus, 1988). Policy formation and making should be an exercise in value adjustment and value creation (Vickers, 1972) in search for mutual agreement of facts and fictions (Klabbers, 1985).

**3. Adversarial**

As explained before, assumptions play an important role in the problem structuring process. However, many policy makers are unaware of the fact that their interpretation of the problem is based on such assumptions, let alone that they realize what the assumptions are that they use to interpret the problematic situation. In other words, *"most policymakers are unaware of the fact that much of their action rests on assumptions and, moreover, they are unaware of the particular set of assumptions they hold"* (Mason & Mitroff, 1981, p. 18).

To increase the policy makers' awareness of their own and others' set of assumptions, and to provide an opportunity to test and compare their assumptions, the policy method should be adversarial. Stated differently, to decide which elements (assumptions or assumed knowledge) to incorporate in the common model of the problem and which to exclude, a dialectical method should be used: in the problem structuring stage, underlying assumptions should be made explicit, differences in position should be questioned, and evidence should be gathered to arrive at some final conclusions.

**4. Managerial mind supporting**

The next implication for the design of policy methods is that they should support the policy maker's understanding of the problem rather than focusing on the behavioural aspects (such as decision-making) of the persons job. In other words, the policy methods should be 'managerial mind supporting' (Mason & Mitroff, 1981, p. 16):

> "The choice of individual courses of action is only a part of the manager's or policy maker's need. More important is the need to achieve insight into the nature of the complexity and to formulate concepts and world views for coping with it. It is the policy maker's thinking process and his or her mind that needs to be supported."

Since participative policy modelling focuses on the problem structuring process, the process where people rethink the way in which they perceive and represent the problem, the managerial mind supporting criterion should not come as a surprise.

**5. Context-dependent or situational**

Finally, it is important that the policy method to be selected is able to take into account the context within which the process of shaping and articulating knowledge is taking place. The purpose of the policy method should not be the production of context-independent knowledge (e.g. a conceptualization that is valid 'for ever', independent of the context, culture, and people who participated in its construction), but an understanding of the situation in the context of the situation itself. *"Knowledge is situated, being in part a product of the activity, context, and culture in which it is developed and used"* (Brown, Collins, and Duguid, 1989). It is because of this that Klabbers (1990) states that: *"most of the problematic situations we encounter like for example, environmental degradation, drugs- and crime prevention, health care, starvation, etc. cannot be resolved by narrowly technical means based on a rationalist conception of reality (...) a historicist approach in which the rationalist conception of reality is embedded, as for example the actor-approach, is more promising and certainly more playful"*. Policy methods that take into account this particular aspect of policy making and complexity, have an eye for the historical nature of the outcome and realize that they are concerned with a process that is irreversible, non-repetitive, and unique in at least one of the factors location, actors, and time (Vickers, 1972).

Now that we have discussed what should be expected from policy methods for ill-structured problems, it is important to consider whether or not participative policy modelling is able to meet the criteria just mentioned.

Before we are in the position to do so, however, a brief description needs to be given of the participative policy modelling method or programme itself. The next section will discuss the main elements of the participative policy modelling method from a historical point of view, that is, starting with a description of its ancestor: the system dynamics modelling approach.

## 2.5 PARTICIPATIVE VERSUS TRADITIONAL POLICY MODELLING

Participative policy modelling builds upon a mathematical modelling technique that is called system dynamics. To understand and appreciate the participative nature of participative policy modelling, we will begin our description of participative policy modelling with a brief discussion of the system dynamics approach and the reasons why a more participative version was developed. See also Vennix (1990), Chapter 1.

### 2.5.1 SYSTEM DYNAMICS MODELLING

System dynamics was founded by Forrester in the late 1950s as Industrial Dynamics (Forrester, 1961). Employing concepts from disciplines such as cybernetics, engineering and management science, a systems approach was developed to tackle problems in industry (for instance problems related to logistics). Later on, system dynamics modelling was used for modelling urban problems. However, it did not become well-known by the public until its application to global problems (Meadows et al., 1982).

In all these applications, the major activity in the systems approach consists of creating a *model* that can be used to examine and understand the dynamics of the problem at hand. The model is built to represent the aspects most important to the problem only and thus serves to simplify the problem to make it more manageable. Moreover, because the system dynamics model to be constructed is a formal one (a mathematical model), computer analyses can be carried out to provide additional insight, in particular in the model's relationship between structure and dynamics. Without such computer analyses, it is difficult to understand, let alone to predict how changes in, say the structure of the model, will affect its behaviour: *"...although the models in the human mind are complex, most people can deal with only three or four variables at a time and do so through only one or two time iterations"* (De Geus, 1988, p. 74). Computer analyses thus can be used to increase one's understanding of the situation as it is (status quo) and to examine the (dynamic) effects of potential policy alternatives - to study the impact policy measures have on the variables that are part of the model.

To differentiate system dynamics models from other modelling approaches (for instance, econometrics) the following four characteristics of system dynamics modelling can be discerned[9]:

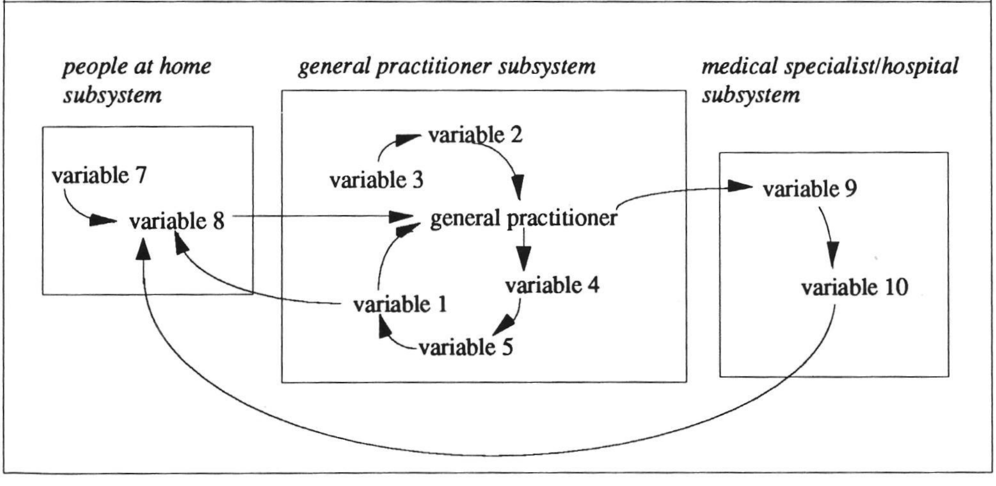
**1. Holistic**

System dynamics modelling considers the systems or problem to be modelled as a whole rather than focusing on one particular aspect of the problem (for instance the economic or psychological aspect). In this sense, holistic refers to the absence of an a priori conceptual framework (regarding a content area) used to interpret and analyze the problematic situation. It also stresses the importance of a methodological point of departure as a consequence of which aspects from various disciplines can be included.

The second way in which system dynamics models are holistic is because they do not focus on the factors that are closely related to the problem area only, but also on factors that are important but somewhat distinct from the core variables of the model. In this case, the term holistic refers to the scope of the model rather than the content.

To illustrate the difference between the two ways in which a system dynamics model can be holistic, look at a hypothetical model of the Dutch health care system (figure 2.6).

Figure 2.6: A simplified model of the Dutch general practitioner part of the health care system




As depicted in the figure, the model concentrates on the variables that are related to the general practitioner. Since the variables that are included may stem from disciplines such as psychology, medicine, economics and sociology, the model can be characterized as multi-disciplinary, and consequently holistic. However, because the model does not only

9 For a description of the differences of system dynamics modelling and other mathematical modelling approaches, see Meadows, 1980; Meadows and Robinson, 1985; Vennix, 1990; Vennix, Smits, and Geurts, 1989.

focus on variables that are close to the general practitioner, it also includes factors that belong to the fields of the medical specialist and/or hospital and people at home (because those factors do affect the general practitioner subsystem as well), its scope is such that the model can be called holistic as well.

**2. Internal structure / feedbackprinciple**

Another, probable the most important characteristic of the system dynamics approach is its emphasis on the internal structure of the system as the cause of its dynamic behaviour. According to the system dynamics approach, it is the structure of the variables, the shape of the network, that accounts for most of the dynamic behaviour of the individual variables (Meadows & Robinson, 1985). The importance of the structure is reflected in the role the notion of two-way causation or feedback (decisions or actions in the system cause changes in the system which consequently might affect future decisions or actions) plays in the construction and analysis of the models. Feedback structures are important because of their capacity to produce certain types of dynamic behaviour of the system. They can bring about dynamic behaviour such as decay, growth, and, instability (see figure 2.7).

Figure 2.7: Three modes of dynamic behavior

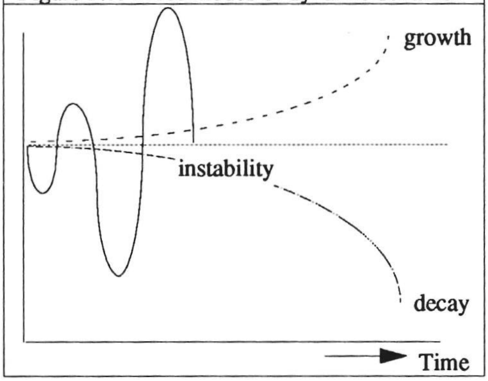
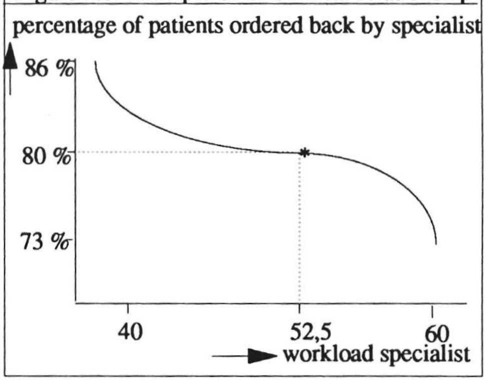



Figure 2.8: Example of a nonlinear relationship




Because of the impact the *shape* of the network has on the model's dynamic behaviour, thinking about the factors that need to be included in the model and the way in which they are related to each other is the most important part of the model-building process. Conceptualization thus forms the heart of the system dynamics modelling process.

**3. Non-linearity**

In contrast to some other mathematical modelling methods, system dynamics does not (for the sake of simplicity) require linear relationships, but also allows for non-linearity. One way to build these nonlinear relationships into the model is to make use of a table function, which can be presented graphically as depicted in figure 2.8.

The figure shows that the relationship between the variables 'percentage of patients that are being ordered back' and 'workload medical specialist' is conceptualized as being *nonlinear*. In the normal situation, when the workload is equal to 52.5 patient per week (per medical specialist), the percentage of patients ordered back is 80. However, changes in the workload result in changes in the percentage of patients ordered back. The reason for this, we assume, is to keep the workload at a fairly constant level. Increasing the number of patients ordered back will lead to an increase in the workload and vice versa. Since the degree to which 'changes in the workload' affect the 'percentage of patients ordered back' varies depending on the value of the workload, the relationship between the two variables can be considered as being non-linear in nature.

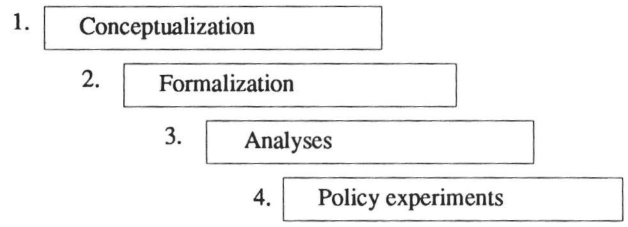
**4. Delay**

The fourth and last characteristic of system dynamics modelling is its ability to take into account delays (e.g. delay of information). An example of such a delay is the time it takes before a general practitioner realises that his or her *structural* workload has changed. It is only, say, after two or three busy weeks that a general practitioner starts to notice that the change in workload is not the result of a nonstructural change such as an influenza epidemic and that it is of a more structural nature.

Delays play an important role in the stabilization or destabilization of the behaviour of a system. If, for instance, decisions are taken too early or a too (due to inappropriate delays), the behaviour may be exacerbated rather than stabilized.

Having discussed some of the core elements of the system dynamics approach, it is time to examine in more detail the **process** by means of which a model is constructed. The description is based on system dynamicists such as Forrester (1961, 1969), Klabbers (1977), Randers (1980), Richardson and Pugh (1981), Roberts, Andersen, Deal, Garet, and Schaffer (1983), and Vennix (1990). They all look upon the model-building process as a process divided into **stages**. Since a brief description of these stages will contribute to our understanding of system dynamics modelling, each of them will be discussed in some detail.

Figure 2.9: Main stages in computer modelling




On the basis of a preliminary definition of the problem and its definition of the system's boundaries, time horizon, and level of aggregation, the process is started by identifying the elements relevant to the problem at hand and the relationships among those elements. The model-building stage in which these activities are carried out is called the ***conceptualization stage***. It results in a conceptual model in which most of the important variables and their interrelationships are visually represented[10]. Usually, a verbal model is added to the conceptual model to describe rather than visualize the most important variables and relationships.

Once the conceptualization has taken place, the conceptual model is translated into mathematical equations and the model's parameters are being quantified. The ***formalization stage*** thus results in a formal description of the model which can be used to create the computer model.

Following the formalization, the model-building process enters the ***analysis stage*** where analyses are carried out to check for logical values of the model variables; to see whether or not the model is capable of producing logically consistent results. Subsequently, the model is validated to assess if and to what degree the model is a valid

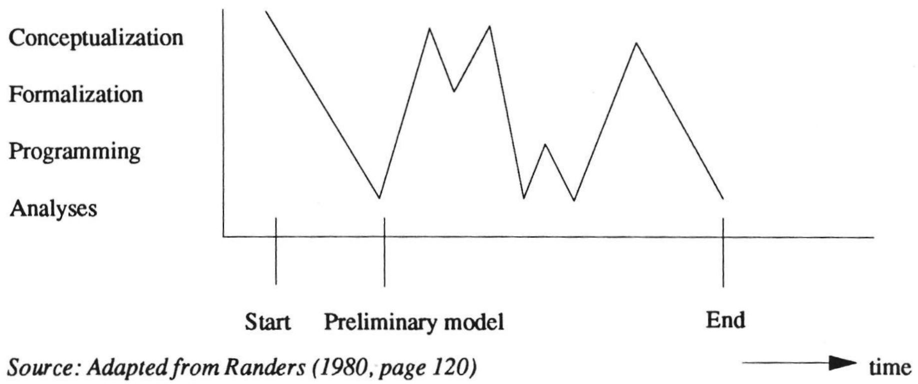
10 For an example of what a conceptual model may look like, see figure 2.6.

representation of the problem at hand. Moreover, sensitivity analyses are carried out to examine the relationship between the structure of the model (shape) and its dynamic behaviour. Based on the knowledge acquired in the analysis stage, model-builders move on to the next and last stage of the model-building process, the stage where policy experiments are carried out.

In the ***policy experiment stage***, policy experiments can be carried out in two different ways. First of all, it is possible to implemented policy alternatives in the model to examine their effects. Secondly, policy experiments can be carried out by using the model to assess under what conditions which policy alternatives can be realized.

Despite the 'linear' presentation of the model-building process (one stage after the other), note that the model-building-process-*in-use* is of a cyclic rather than a linear nature, as depicted by Randers (1980, p. 120):

Figure 2.10: The model building process




*Source: Adapted from Randers (1980, page 120)*

The horizonal axis shows the elapsed time in the system dynamics modelling process. The vertical axis distinguishes between the four major stages in the modelling process where the cyclic character is predominant. Note that the policy experiment stage has been excluded to make room for the programming stage. The programming stage can be considered as a substage in between the formalization and analysis stage. It covers the transformation of the formal model into a computer model.

The figure shows that in the beginning of the project, one relatively quickly moves from one stage to another; from the first conceptualization to the first analyses. However, during the first analyses, it is often discovered that adaptations are needed to both the structure of the model and its parameters. As a consequence, model builders have to return to the conceptualization stage and rethink the original model. As depicted in the figure, this is followed by some going back and forth between various stages until the results are satisfying.

Having described the system dynamics modelling process in some detail, we will briefly discuss how well these system dynamics projects succeeded in having an impact on the policy making process.

## 2.5.2 EVALUATION OF (TRADITIONAL) SYSTEM DYNAMICS MODELLING

Following Vennix (1990)[11], three three categories of problems that many traditional system dynamics modelling have encountered can be distinguished: problems related to the *models* themselves, problems related to the model-building *procedure*, and finally, problems related to the *use* of system dynamics modelling.

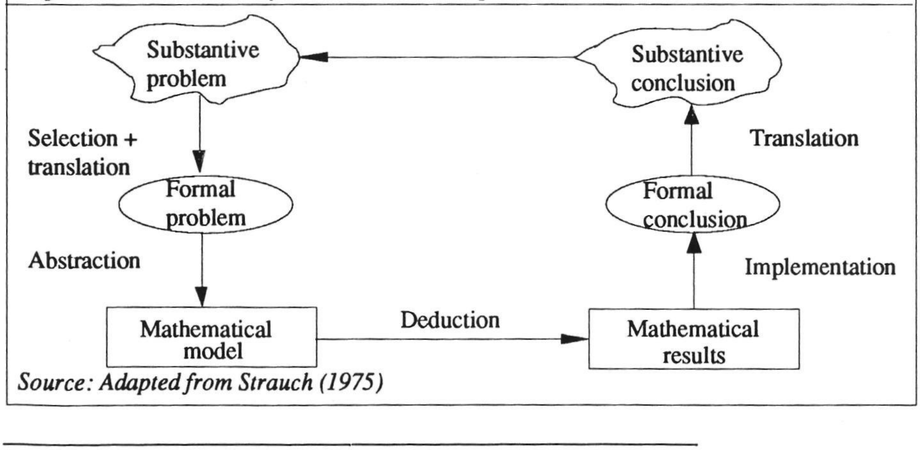
Model

Regarding the first category, it was found that in the early days, model-builders were too ambitious; the models they built were too comprehensive (they almost tried to model the whole world rather than a particular policy problem), and because of poor documentation, these models were almost inaccessible for outsiders.

Procedure

The major problem in relationship to the model-building procedure concerns the lack of client involvement in the traditional system dynamics modelling approach. Differences in orientation between modellers and policy makers, as for instance expressed in the Two-Communities Theory (Caplan, 1979), show that without client participation model-building projects have only limited impact. Without client participation, it will be difficult, if not impossible to communicate the outcomes of the modelling project to the world of policy making. Or as Weil (1980) puts it: "*... we have learned that we cannot just deliver our recommendations and ride off in the sunset, in what Frohmann cites as the 'Lone Ranger' syndrome.*" One reason for this is that in modelling an ill-structured problem, one has to translate the 'problematic situation' (the problem felt by the policy makers) into a formal model, carry out some analyses, and translate these formal conclusions back into the world of the policy makers (see figure 2.11). If policy makers do not participate in the process of formalizing the substantive problem into a formal model and deformalizing the formal conclusions into a substantive conclusion (that is, in terms of their own 'material' framework), it is unlikely that the results are being understood, accepted, and used.

Figure 2.11: The model cycle for ill-structured problems




*Source: Adapted from Strauch (1975)*

---

11 For an excellent overview of the findings of evaluation of the impact of computer policy models on (public) policy making, see Vennix, 1990, p. 35, table 2.1.

Also, by having the policy makers participate in the model building process, it is more likely that their own assumptions are being incorporated into the model as a result of which it will be easier for them to understand and trust the model: *"If a planner walks into a room with a model on his computer that he has made up himself, the chances are slim that his audience will recognize this particular microworld"* (De Geus, 1988, p. 73).

Moreover, as described by Meadows & Robinson (1985, p. 370), tackling ill-structured problems requires participation for complex policy problems generally *"...span more than one person's expertise"*. Having those who are involved in the problem participate, allows for the elicitation of expertise that is required to develop a realistic representation of the problem at hand.

Besides, participation is important because much of the learning about the problem is taking place in the process of building the model (Toval & Flores, 1987; Vennix, 1990). Since research reports do not include a description of the process by means of which the final model was arrived at, knowledge based on the process of going back and forth the various stages, cannot be acquired by reading the final report only.

Without participation integration of the different perspectives into a shared understanding is difficult to realize. By having the people build a model, they create their very own frame of reference. And based on this, they can discuss policy alternatives and decide on what policy actions to take.

Finally, as stressed by Klabbers (1990), participation is important in that it provides the opportunity to use (and (re)construct) system dynamics models in an non-trivial rather then trivial way, that is, it allows for an approach in which the model (or knowledge) of the situation is constructed from a participant/insider point of view, rather than from an observer/outsider point of view. As a result, the notion of multiple reality, that is, the differences in perspective among the participants, is taken into account and the construction of the model is taking place in close relationship to its historical context in terms of actors, time and location (Vickers, 1972).

## Use

The third category of problems are those related to the way in which system dynamics models used to be employed. Rather than using them as a communication device and a starting point for policy, or as a device to generate hypotheses and policy-oriented theories, system dynamics was used to arrive at accurate predictions. Applied to ill-structured problems, this proved to be an unachievable goal. The reason for this is that regarding ill-structured problems we do not have the knowledge (and data) to build a model on the basis of which we can predict say, the exact number of hospital beds that are required in 2010. Instead of using system dynamics modelling as a predictive device, it was recommended to use it as a tool to acquire insight in the problem at hand (to find out how much we know or do not know about the complex problem at hand), or to use it as a communication or integration device (communicating and integrating the differences in perspective: make people aware of differences in perspective and have them somehow integrate them in a 'workable' unified frame of reference).

Summarizing the evaluation of traditional system dynamics modelling, it has been shown that system dynamics has a lot to offer as far as understanding and managing the complexity and dynamics of ill-structured problems is concerned. However, three adjustments should be made to enhance its impact on the policy making process: increase the participation of the policy makers, reduce the scope of the model (model a problem), and finally, use the system dynamics modelling to assist in the conceptualization of the problem, that is to define the problem and to generate policy alternatives rather than using it for predictive purposes.

### 2.5.3 PARTICIPATIVE POLICY MODELLING

Now that we have discussed why and where the traditional system dynamics approach had to be adjusted to enhance its impact on the policy making process in general and the conceptualization stage in particular, a description can be given of the result of these adjustments, that is, the ***participative policy modelling*** process.

Generally speaking, participative policy modelling as a method is based on system dynamics modelling. It aims to contribute to the *problem structuring stage(s)* of the policy making process by *involving the policymakers* (or stakeholders) in the process of identifying the factors that are related to one *particular policy problem* in order to arrive at a conceptual and computer model of the problem at hand. As such, its objective is to enrich the individual participant's conceptualization and to arrive at a so-called shared understanding.

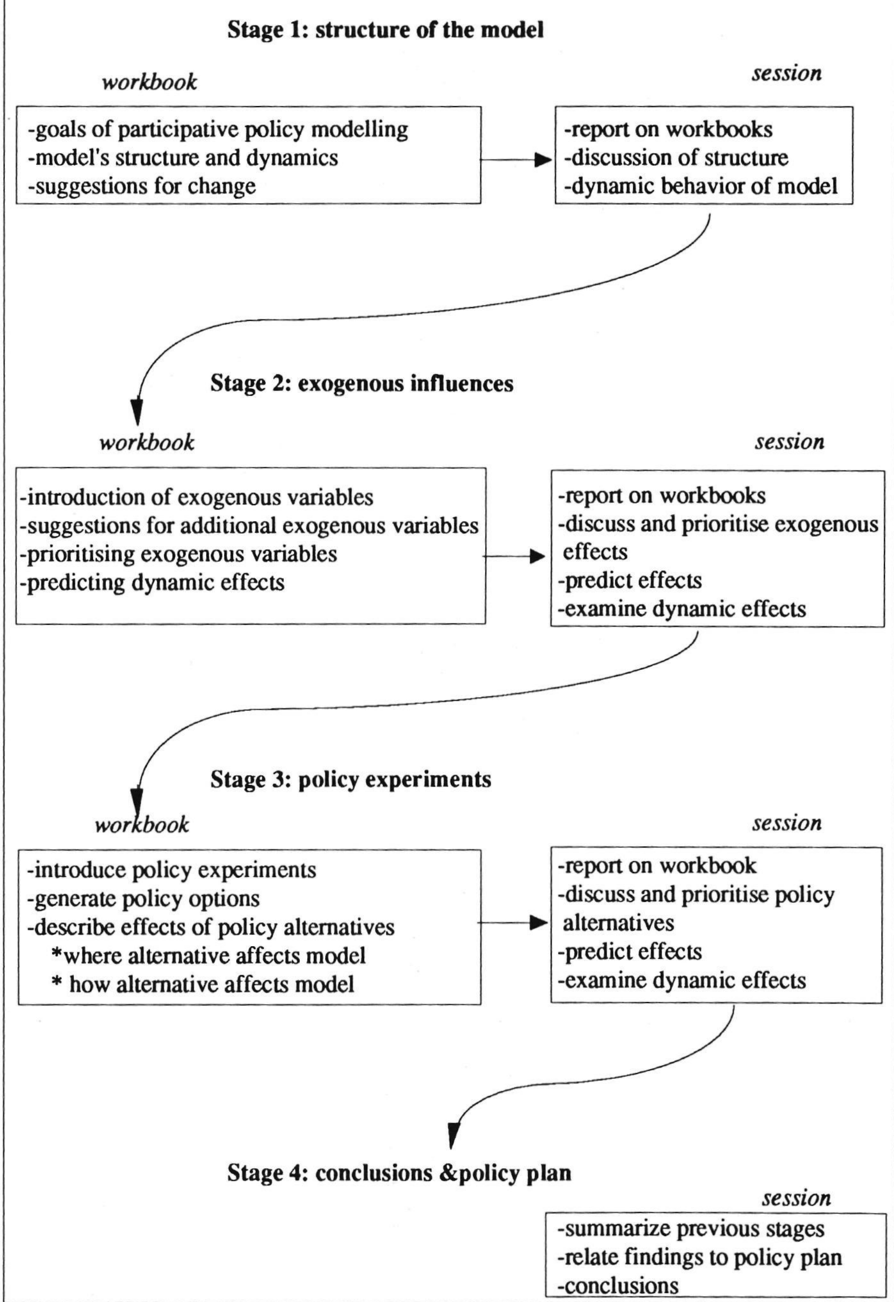
To meet these objectives, participation of the policy makers in the process of building and analyzing the model is required. However, in practice, full participation, that is involvement in all four major stages of the modelling process, is difficult to realize. The first reason why policy makers hardly ever participate in all stages is because of the amount of time that is required for full participation. Policy makers tend to be very busy people and simply don't have (or take) the time required to participate in all stages. Moreover, participation in the formalization and analysis stages is difficult because of the technical skills that are being called upon (for instance translating the model in mathematical equations and programming the model in a computer language) in those stages. Although user-friendly software packages such as 'Stella' and 'I Think' have been developed to overcome these problems (in particular to overcome the difficulties people have in formalizing the model), analyses (e.g. sensitivity analyses) still take a lot of time and are certainly not the policy makers' most preferred model-building activity. They rather participate in the conceptualization and policy experiments stages than in what they consider to be the more tedious stages of the model-building process - the formalization and analysis stages, a problem also encountered by Vennix (1990).

Taking into account these restrictions, it was decided to include as much participation as possible, in particular in the conceptualization and policy experiments stages, and to have model-building experts carry out the programming and analyses tasks. Obviously, this restricts the participative character of participative policy modelling, but it was felt that in order to meet the objectives of the participative policy modelling (to enable the participants to *exchange their ideas (or expertise)* about the problem at hand, to arrive at a *commonly agreed upon model* (definition, description) of the problem, and to generate ideas on *how the problem should be dealt with*), participation in the conceptualization and policy experiments stages would be more important than participation in the other two stages.

Hence, the term 'participative' does not refer to the ideal program in which participants participate in all phases of the model-building process to the same degree. Participation is often restricted to one or two phases in particular. In the present study, participation is restricted to making adjustments to a preliminary model rather than building the model from scratch. For this kind of participative policy modelling the terms computer-assisted learning environment and flight-simulator are regularly used as well (Morecroft, 1988; Kim, 1989; Senge, 1989; Vennix, 1990). In the present study, the term 'participative policy modelling' is used to refer to the model-building process because participants are given the opportunity to construct a model of their own. However, to speed up the building process, a so-called preliminary model is used. Note that this preliminary model is introduced to the participants as being preliminary - participants are told not to hesitate and change as much as they want of the model to make it in line with their perception of the system.

Taking into account the above considerations, the actual participative policy modelling program carried out in the present study has been designed as follows (Vennix, Verburgh, Gubbels and Post, 1990):

Figure 2.12: Detailed design of the participative policy modelling sessions of the present study



The figure shows that most of the ***conceptual activities*** (identifying the most important variables and the way in which they are related to each other) are taking place in the first two stages.

In stage one, participants construct a model representing the most important 'endogenous' factors. Endogenous factors are factors that both affect and are being affected by variables related to the problem. It is important to focus on the network of endogenous factors to start with because, as explained before, it is the structure of this network that determines the model's dynamic behaviour. To arrive at a causal model of the endogenous factors, participants select the factors that seem most important to the problem and discuss how they are related. Once the causal model is constructed, participants are invited to examine the model's dynamic behaviour (i.e. examine the changes in the values of the variables over time).

In the second stage, people concentrate on exogenous factors that may have an important effect on the behaviour of the model created in the first stage. Exogenous factors are factors that have an impact on the causal network developed in stage one, but are themselves not affected by the network. Obviously, one has to take into account the model's time-horizon, for it may well be that by increasing the time-horizon (say from 10 to 100 years) a particular exogenous variable suddenly is becoming affected by the behaviour of the model. In other words, feedback mechanisms that were not visible on a 10 year's time-scale may become visible when the time-scale is increased to 100 year. An example of an exogenous variable related to the present study's model of the Dutch health care system is the variable 'ageing population'. Having identified the exogenous factors that may impact the costs or volume of the health care system, participants decide which of the exogenous factors will be most important (or interesting) and consequently should be examined in more detail.

As part of this closer examination, participants are asked to predict what the effects of the exogenous variable will be on the behaviour of the model (e.g. how much will the workload of the general practitioner increase assuming that the ageing population is affecting the volume as agreed upon by the participants) and compare their expectations to the dynamic behaviour calculated by the simulation model (based on the assumed relationship between impact of the exogenous variable on the model). If a discrepancy exists, and this is often the case for it is difficult to oversee the dynamic consequences of a set of singular assumptions like 'if the average age of the population increases by 1 per cent, the number of patients consulting their general practitioner for the first time will increase by .05 percent' and 'if the average age of the population increases by 1 per cent, the number of patients referred back from the medical specialist to the general practitioner will increase by .3 per cent', participants are encouraged to discuss whether and how adjustments need to be made to their own individual or collective (shared) representation of the model (including the way in which the relationships between the exogenous variable and the model was being conceptualized). In other words, if there is a discrepancy, people either change their individual expectations, that is, change their individual mental model (Waern, 1990), or adjust the external model, that is, the system dynamics model developed in stages one and two. Regarding the adjustments to the external model, changes can be made with respect to the internal causal network or the way in which the exogenous factor is related to the endogenous model.

In the third stage of the participative policy modelling program, ***policy experiments*** are being focused upon. Based on the model developed and discussed in stages one and two, participants consider a number of policy experiments and decide which one to examine in more detail first. As part of this in-depth examination of one particular policy alternative, participants discuss and predict the policy alternative's effect on the model's behaviour and contrast their expectations to the dynamic behaviour generated by the computer model. Note that the model is used to *"produce information about the consequences of adopting an alternative"* (Quade, 1989, p. 140) for it is only *"in rare instances that we can predict with near certainty or assign high confidence to the results that come out of the model."* (Quade, 1989, pp. 140-141).

In stage 4, the preceding activities are wrapped up and related to the already existing (if any) policy plan. People are invited to discuss how the model-building procedure is affecting the plan. The fourth stage is mainly used as a reporting and dissemination stage, geared to the needs of the organization.

Note that the program distinguishes between two kinds of activities: preparatory ones where participants read and fill out a workbook, and sessions in which the actual participative policy modelling is taking place. To speed up the model-building process, participants are asked to read workbooks and complete exercises to prepare for the next session. Once these workbooks are handed in, they are analyzed by the expert model builders/facilitators to complete their preparation of the next session. Suggested adjustments to the preliminary model (that is, the model developed by the expert model builders together with one or more experts in the particular content area, and which serves as a catalyst of the building process, again to speed up the process) for instance, are used to prepare for potential adjustments to be made to the computer program during the session and to structure the program of the session.

Not visible in figure 2.12, but clearly an important characteristic of the program is its variety in instructional methods. Usually, participative policy modelling sessions contain activities that are carried out in the lecture mode, small-group activities mode, and plenary discussion mode.

Generally speaking, sessions start with the lecture mode to report the results of the completed workbooks. This is followed by a small-group activity which really forms the heart of the program. During these small-group activities participants work in small groups in break-out rooms where they discuss particular topics using pre-structured materials such as worksheets. At the end of the small-group activity, one of the group members is assigned as a spokesperson and prepares shortly to present the group's findings, solutions or conclusions to the rest of the participants.

After a short coffee break, allowing the facilitators to briefly discuss and evaluate the small-group activity and to prepare for the plenary discussion, each of the spokespersons presents the findings of his or her group. In the following plenary discussion, the rest of the participants can respond to the presentation. The plenary discussion results in a list of adjustments to the preliminary model that are agreed upon by the participants, and, if there are any, a list of adjustments that are not agreed upon by the participants.

Figure 2.13: Participative policy modelling session

| | mode of instructing | time scheduled |
|---|---|---|
| 1. report of the workbooks | lecture mode | 15 minutes |
| 2. small-group activity | small-group discussion mode | |
| * discussion | | 60 minutes |
| * computer analyses | | 30 minutes |
| 3. coffee-break | | 15 minutes |
| 4. plenary presentation and discussion | discussion mode | 60 minutes |
| | | total time: 3 hours |

Having explained the participative policy modelling program in some detail, it is time to return to the question of why participative policy modelling as a policy method is particularly suited to ill-structured problems. Recall that to answer this question, a description was given of the features of ill-structured problems. Based on this description, five requirements were derived that methods for ill-structured problems should meet. In order to be suited to ill-structured problems, it was stated that policy methods should not only be participative, integrative, adversarial, and managerial mind supporting (Mason & Mitroff, 1981), but also context-dependent (Brown, Collins, and Duguid, 1989; Kalff, 1989; Klabbers, 1990; Prawat, 1989).

As far as the requirement of participativeness is concerned, it has been shown that participative policy modelling is **participative** in that it involves the policy makers in the construction of the model (Graham et al., 1989; Kim, 1989; Morecroft, 1988). Policy makers are encouraged to draw upon their own expertise to develop a commonly agreed upon representation of the problem. The model-building experts only serve as facilitators of the process and do not interfere in the discussions that take place among the experts.

However, participants do more than just reveal their own point of view (mental model) in the participative policy modelling process. They also test and compare their assumptions (mental maps) by discussing their point of view with the other participants and relating their assumptions to the outcomes of the computer programme. In other words, participative policy modelling is **adversarial** in that it not only increases the participants' awareness of the fact that they have assumptions, but also what these assumptions are and whether or not they are in agreement with the assumptions that are held by the other participants.

Participative policy modelling thus contributes to a divergence in the way in which the problem is being conceptualized. As a consequence, it decreases the likelihood of a premature closure or jumping at conclusions. However, it is important that the process of questioning the assumptions underlying their conceptualization does not lead to *"down blind and narrow alleys"* (Churchman, 1971, p. 176). Diversification needs to be followed by a process of **integration**. Participative policy modelling is taking the integrative aspect of dealing with ill-structured problems seriously in that it explicitly aims at arriving at a shared representation of the problem by having the participants build an agreed upon conceptual and computer model.

Regarding the **managerial mind supporting** requirement, it should be clear by now that the major objective of participative policy modelling is to enrich the way in which policy makers look upon the problem. Its concern for the conceptualization stage of the model-building process, and its focus on the early stages of the policy making process (where thinking about the problem and how to approach it are predominant) manifest the managerial mind supporting orientation of the participative policy modelling. Participative policy modelling's primarily concern is not doing something about the problem (in the long run, it obviously is concerned with policy action), but getting to *know* what to do.

Finally, the participative policy modelling's participative character enables the method to meet the fifth requirement, according to which the participant or insider point of view should be acquired to be able to really understand the complex process in its historical context. By having the people that work in that field, for whom the problematic situation is a problem and who have knowledge of the problem's **context** and history (in terms of actors, time, and location), participate in the construction in the model, a model is created that is not separated from the context in which it is used and from which it derives its meaning: participants will become aware of the fact that the *"behaviour of a system dynamics model depends on its initial conditions, and on the mix of adjustments in parameters"* (Klabbers, 1990). It is for reasons like this that Klabbers (1985, 1990) and Kalff (1989) claim that participative (or interactive) policy modelling can be used to both enrich the individual participant's point of view, and bring about a joint perception of the problem at hand: *"The characteristics of simulation model building by management teams meet many of the new requirements for strategic decision making"* (Kalff, 1989, p.60).

Following the description of the major characteristics of participative policy modelling in relationship to the features of both policy making and ill-structured problems, one additional requirement that policy making in general and the use of participative policy modelling in particular must meet, needs to be discussed. It is only now that an overview of the participative policy modelling approach has been given, that we feel its discussion is appropriate. The requirement concerns the relationship between participative policy modelling and its subject in terms of stability (over time).

In order to be able to apply the participative policy modelling approach, the system under consideration should have *"endurance through time"* (Vickers, 1972). The model-to-be should be concerned with conditions and relationships that are maintained through time (at least for the time of the participative policy modelling program) to avoid the previously discussed fallacy of representing/solving the wrong problem - the so-called error of the third kind (Raiffa, 1968). If the system does not have a particular degree of stability, approaches different from the participative policy modelling one should be employed.[12] Hence the present study needs to show that, in spite of the turbulent changes in the Dutch Health Care system, a reasonably stable representation can be arrived at to ensure that the conclusions based on this representation still hold within the very same definition of the problem. The stability of the system will be dealt with in Chapter 5 in which a description of the Dutch Health Care system (and its model) will be given.

Having discussed the reasons why participative policy modelling seems to be applicable when dealing with ill-structured policy problems in the context of policy making, and having explained the present study's evaluative concerns, an overview of the state of the art, as far as evaluating the effects of participative policy modelling is concerned, will be given next.

## 2.6 EVALUATING PARTICIPATIVE POLICY MODELLING: THE STATE OF THE ART

Having described in some detail the features of participative policy modelling[13], account needs to be taken of the evaluations that have been carried out so far. This to appreciate the distinctive features of the present study's attempt to assess the impact of participative policy modelling on the policy making process in general and the policy makers' conceptualization of the problem in particular.

Except for Hart (1985), Kaplan, Lombardo, and Mazique (1985), Klabbers (1972), and Vennix (1990), most of the above mentioned applications do not include a formal evaluation. As mentioned by Graham et al. (1989, p. 325), *"work on measuring methods is progressing in several projects. As mentioned above, the experiments described in Moissis (1989), Gould (1989b) and Bakken (1989) are under way at MIT."* In other words, the importance of evaluative studies is being recognized, but so far hardly any empirical evidence for the effectiveness of participative policy modelling has been given. Or as Kim (1989, p. 332) puts it:

> "To date, we have no empirical evidence for any of the variables mentioned above nor on the connection of "better management thinking" to bottom line numbers... Anecdotal evidence[14] suggests, however, that the LL [learning laboratory] concept

---

[12] Klabbers (1985, p. 141) developed a taxonomy for policy methods on the basis of time horizon and degree of accuracy of description of social and societal systems.

[13] For applications of modelling that are, some more than others, based on system dynamics modelling, see for instance Hart et al. (1985), Hart (1985), Kaplan, Lombardo and Mazique (1985), Klabbers (1976), Morecroft (1988), Morecroft, Lane, and Viita (1989), Richardson, Vennix, Andersen, Rohrbaugh, and Wallace (1989), Vennix, Gubbels, Post, and Poppen (1988). Recently, people have started to use terms like 'computer-assisted learning environment' and learning laboratories (Kim, 1989; Senge, 1989; Senge & Sterman, 1990; Vennix, 1990; Vennix and Scheper 1990) to refer to participative versions of system dynamics modelling.

[14] See for instance Morecroft, 1989, p. 175: "The commercial manager said that he had learned a lot about the factory from the project and had been made aware of the critical importance of distributor loss (and the factors that influence loss) in limiting the growth potential of the business".

has great potential for helping managers reassess the way they think about their business."

Hart (1985), and Kaplan et al. (1985) by contrast, have included some empirical evaluation. In all three studies, participants were asked to fill out a questionnaire on what they had accomplished and the process by which this had occurred. The three studies focus on the *perceived* effectiveness of both the outcomes and the process by means of which the outcomes are brought about. The Hart (1985) study also reports the use of additional client and staff interviews for comparative analysis. However, because they were based on particular projects, they were not included in the reported evaluation.

Although the studies mentioned above are an important step forward in the process of empirically evaluating the process and products of participative policy modelling, it was felt that a more objective approach (i.e. an approach which does not completely rely on self-ratings, measuring the participants' perceptions of the process and outcomes) would be more appropriate. Hence, Vennix (1990) employed an experimental design (pretest-posttest control group design) in which not only a questionnaire was used to examine the participants' perceptions of the process and outcomes, *but* participants were also asked to both write a policy note and fill out a multiple-choice test, measuring their knowledge of the problem at hand before and after completion of the course, to evaluate the conceptual impact of participation more objectively. To summarize his approach:

Figure 2.14: Overview of the Vennix 1990 study

| | sessions | | | | | |
|---|---|---|---|---|---|---|
| pretest | 1 | 2 | 3 | 4 | 5 | posttest |
| $X_1$<br>* policy note<br>* multiple-choice test | questionnaire | questionnaire | questionnaire | questionnaire | questionnaire | $X_2$<br>* policy note<br>* multiple-choice test |

*Population*: students
*Subject*: social security system of the Netherlands, hence used an econometric model
*Design*: pretest-posttest design
*Individual* changes in conceptualization
Quality defined in terms of *policy theory*

As depicted in the figure, the participants were all students, mainly from the Social Sciences Department. They were assigned randomly to an experimental and a control group. The experimental group took part in participative policy modelling sessions (called 'computer-based learning environments by Vennix), while the control group was being taught in the traditional lecture mode.

The content of the course focused on the structure and dynamics of the *Social Security System* of the Netherlands. Because a system dynamics model of the Dutch Social Security Model was not available, an econometric one was being used (cf. Douben, 1987). Since there are considerable differences between system dynamics modelling and *econometric* modelling[15], this may have affected the study's findings. It is difficult to

15 For an overview of these differences see for instance Meadows, 1980; Vennix et al., 1989; Vennix, 1990.

expect participants to include non-linearity in their conceptualization when the model used to show them the importance of non-linearity only includes linear relationships. Moreover, we feel that the use of an econometric model has severely limited the degree to which his participants were able to participate in the construction of the model. Whether or not participants had the opportunity to make any changes to the model (in fact, participants could carry out analyses but not change the model's structure), one can doubt whether economics is the best topic to encourage participants to think about and challenge the knowledge that is being presented. This because economics is a relatively well-developed theoretical framework. In other words, it is believed that the Social Science students may have had difficulties challenging an econometric model because economics is being perceived as sound in theory (certainly if compared to the knowledge people have about ill-structured problems). Our objections to the Vennix application thus concern both the content of the application and the kind of model that is being used.

Another feature of this study concerns the level at which the conceptual impact of the participative policy modelling is being assessed. In the study, *individual participants* are being focused upon with respect to both the level of measurement and the level of analysis. Based on the individual participants' pre- and posttests scores, the degree to which intra-individual changes have taken place is being assessed. The study briefly discusses the integration of the various perspectives on the problem under the heading of 'organizational learning', but does not include the inter-individual changes (or organizational changes) in the analyses and results.

Moreover, the criteria which are used to assess the degree to which the participants change their representation of the problem (applied to the policy notes they had to write) were derived from Dutch literature on policy theories (cf. Hoogerwerf, 1984; Kraan-Jetten, 1986; Leeuw, 1983, 1986; Ringeling, 1985). According to this literature, policy theory can be defined as *"...the total set of assumptions underlying a (specific) policy"* (Hoogerwerf, 1984, p. 594). Assuming that the assumptions policy makers have regarding a particular problem is reflecting the way in which the problem is conceptualized, Vennix decided to use changes in the policy theory to assess changes in the participants' perception of the problem. Based on the literature on policy theory, three major criteria were derived to assess the quality of the policy theory (or conceptualization): an epistemological one (about validity and precision of the policy theory), an implementory one (dealing with policy variables, differentiation and integration) and a strategic one (including societal conditions and time factors). Using Axelrod's (1976) coding procedure, pre- and posttest policy notes were scored on these criteria and a comparison was made between the post- and pretest scores[16].

Although the policy theory approach is an interesting one, it was felt that a different framework should be used to arrive at criteria to assess the impact of participative policy modelling on the conceptualization of the problem. The major reason for this is not that we disagree with Vennix's assumption that changes in the way in which people look upon a problem are reflected in the people's policy theories (in fact we agree), or that the criteria that come from the world of 'Public Administration' are not useful to assess the quality of the policy theories (they seem promising). The major reason is that at this stage of empirical research we believe the most important research question is *not* the assessment of quality in terms of 'inclusion of all the aspects (strategic, implemental and epistemological) that make policy theories useful in the area of Public Administration', but the assessment of quality in terms of the participative policy modelling program itself, that is, in terms of the system dynamics thinking that is being offered in the participative policy modelling program. Participative policy modelling primarily aims to bring about a system dynamics view (or model) of the problem at hand. As a consequence, the effects participation has on the way in which the problem is being looked upon should be assessed in terms of the program offered to the participants (the

16 For a detailed description of the final indicators of the quality of the policy theory, see Vennix, 1990, table 7.5, p. 155.

change objectives that are directly related to the program offered) rather than in terms of some criteria indirectly arrived at. This is not to say that participative policy modelling does not aim to contribute to the 'Public Administration'-quality as well (comparing the Vennix criteria with the criteria used in the present study will show that there is indeed some overlap between the way in which 'quality' is operationalized in both studies). By contrast, the present chapter's main objective was to show that participative policy modelling indeed can be of assistance when dealing with ill-structured problems in the context of policy making. It thus means that at this stage of empirical research (with only little experience available regarding the use of the cognitive mapping procedure, the absence of prior real-life evaluative studies, and only a few significant changes found in the exploratory Vennix (1990) study) together with the fact that the 'Public Administration'-criteria are still under discussion (Van Doorn, 1985; Hoogerwerf, 1989), a more modest approach, concentrating on the method's basic objectives seems more appropriate.

The major differences between the study conducted by Vennix (1990) and the present one can be summarized as follows:

Figure 2.15: Vennix (1990) versus the present study

| | *Vennix* | *present study* |
|---|---|---|
| population | students | policy makers / staff personnel |
| topic | social security system | health care system |
| model | econometric model | system dynamics model |
| research design | pretest-posttest control group | one-group pretest-posttest design |
| pre- & posttest measures | policy note & multiple choice test | policy note |
| level of analysis | individual | individual and inter-individual |
| source for criteria | policy theory | system dynamics |
| type | laboratory study | field study |

The differences regarding the *source for criteria* and *model* have already been discussed and therefore require no further explanation. Moreover, the differences in *population, research design,* and *criteria* that follow from the selection of system dynamics as a source, will be described in more detail in the chapters to come. As a result, they will not be discussed here as well. The differences that need to be examined in more detail, concern the pre- and posttest measures and the level of analysis.

Regarding the difference in *pre- and posttest measures,* note that the present study does not include a multiple-choice test to assess changes in the knowledge people have regarding the policy problem at hand. The reason why a multiple-choice test was not included was because our participants were *not* dealing with a reasonably well-structured problem (recall, a problem where most of the variables and relationships between these variables are known and considered as being 'true'). The problem they were facing was an ill-structured one. Since ill-structured problems lack generally accepted knowledge, a multiple-choice test distinguishing between right and wrong answers cannot be developed. If it could, the problem would not be ill-structured.

Concerning the *level of analysis,* it was decided to include changes in both the individual representations of the problem and the way in which the individual conceptualizations were related to each other to examine whether or not participative policy modelling is capable of bringing about a kind of 'shared understanding'. The importance of a common frame of reference has been stressed by many (e.g. Klabbers, 1985, 1990; Meadows & Robinson, 1985; Quade, 1989; Vennix, 1990). Hence, it was decided to include this aspect of policy making in our analyses.

The present study thus concerns the evaluation of the effects of participative policy modelling on the way in which individual participants as individuals and as a group conceptualize a particular policy problem, that is, the concepts and relationships they use to refer to the problem at hand. To assess these effects, criteria that stem from the world of system dynamics will be used. In anticipation of their description in Chapter 3, note that a distinction can be made between the domain-specific content of the conceptualization and its content-independent format. Evaluating the effects of participative policy modelling without elaborating upon this distinction would certainly result in an assessment of the method's potential to affect the way in which people conceptualize the problem. However, in the context of ill-structured problems, the question has risen as to whether domain-specific knowledge is as important as the more domain-independent, cognitive strategies (Breuer & Kummer, 1990; Breuer, Tennyson, Lippert, 1989; Tennyson, Thurlow and Breuer, 1988) or strategic problems solving skills (Prawat, 1989) that people (can) use to organize and structure the domain-specific content of their conceptualization in an educational context.

Following Klabbers (1990), it was decided to apply the distinction between domain-specific and strategic knowledge to the world of policy making and policy making support.

The reason why it was decided to take this distinction into account when examining the effects of participative policy modelling, is first of all, because the changes in the knowledge base (Prawat, 1989), or domain-specific knowledge, required to deal with the ill-structuredness of the problem (note that as a result of the ill-structuredness no appropriate domain-specific schema can be activated, and the existing knowledge base needs to be elaborated), need to be guided by a particular problem solving framework or problem-solving strategy to determine what elements to include and exclude and how to (re)structure them in the attempt to solve that particular problem:

> Solche Prozesse [higher-order strategies] ermöglichen einer Person,ihr Wissen "umzustrukturieren", so dass eine Situation analysiert, eine Vorstellung von der Situation entwickelt, spezifische Ziele zur Bewältigung der Situation festgelegt und mögliche Lösungen dazu erarbeitet werden können." (Breuer, Tennyson, and Lippert, 1989, p. 1)

Moreover, note that when the existing knowledge base (domain-specific knowledge) is not sufficient to tackle the problem (as is never the case when dealing with ill-structured problems) not only new knowledge needs to be created, but also new criteria need to be applied to evaluate the selection and organization of this knowledge:

> "For example, if a highly trained classical musician is asked to perform jazz, it may require the creation of <u>new knowledge</u> about jazz idioms.(...). It will also require the development of <u>new criteria</u> for determining the quality of the performance based on criteria held by the jazz community." (Tennyson, Thurlow and Breuer, 1988, pp. 157-158) [underlined by the author].

The distinction between domain-specific and strategic knowledge also holds for the present study in which the effects of participative policy modelling on the way in which individual participants look upon the problem is being evaluated. This means that not only a change in or elaboration of the existing domain-specific knowledge (i.e. knowledge of the Dutch Health Care system) is being focused upon, but also account is taken of the strategic knowledge (strategic, because the outcomes of the process, the input for the policy making process, is to a large extent determined by the general problem-solving skills adopted to structure and organize the domain-specific knowledge) incorporated in the conceptualization of the problem. Since the present study aims at having the participants adopt a system dynamics approach in modelling ill-structured policy problems, the strategic knowledge that is being sought for is system dynamics knowledge (in action[17] [18]), that is, use (and thus knowledge) of system dynamics in selecting and organizing the domain-specific elements that are needed to solve the problem. As a consequence, the evaluation of the conceptualization not only concentrates on the changes in content, the domain-specific configuration that is required for the particular situation, but also assesses the quality of that content in terms of system dynamics thinking; the degree to which the participants have been able to adopt (and apply) the principles of system dynamics.

Another reason why the distinction between domain-specific and strategic knowledge is of importance to the present study as a framework to interpret the criteria that stem from the system dynamics approach (are they all strategic, or are some of them domain-specific as well?), is because of the fact that participative policy modelling not only aims to bring about knowledge that is only applicable in one particular situation, in relationship to only one specific ill-structured policy problem. Recall the close relationship between the problem and the model-to-be-constructed and one immediately sees the limited usefulness of the domain-specific aspects of the conceptualization in terms of actors, time, and location. This is not to say that the creation of particular domain-specific schemata is not of importance, for one of the most important objectives of the participative policy modelling method is to contribute to the process of constructing a conceptualization based on domain-specific knowledge of a particular policy problem (particular because of its ill-structuredness and historical uniqueness). However, it stresses the fact that when confronted with another policy problem, for which no appropriate schemata exist, more general problem-solving strategies must be employed (Gick, 1986) - strategies such as system dynamics, that can be applied regardless of domain.

Now that a description has been given of the two aspects of a conceptualization that can be discerned - domain-specific and strategic knowledge - attention can be drawn to the question of how these two aspects are related to each other, in particular in the context of change. For instance, it may be that people increase their domain-specific knowledge, but fail to adopt a system dynamics perspective. As to whether the reverse is possible is more difficult to answer. From a theoretical point of view, the adoption of a system dynamics point of view may result in a decrease (rather than an increase) in the number of domain-specific elements needed to explain the situation, for system dynamics aims to reduce the problem's complexity by concentrating on the most important variables only. Consequently, although (empirical) evidence is lacking, it seems as if the two aspects of a conceptualization of a problem can be changed independently. As to whether a relationship between the two can be established in the present study will be discussed in the chapter in which the outcomes of the present study will be discussed. Note that, from a theoretical point of view, domain-specific knowledge is viewed as, to some extent, conditional to strategic knowledge:

> "Training in problem solving cannot compensate for ignorance in subject matter. Experts in a domain differ from novices not by general problem solving strategies,

---

[17] By knowledge in action (or knowing in action) is meant that the knowledge of the underlying system dynamics framework should not be acquired and considered as 'theoretical' knowledge, isolated from the context in which it is being used. It is not our objective to have the participants adopt isolated context-independent knowledge (neither domain-specific, nor strategic), but rather enculturate them into authentic practices through activity, and social interaction in a way similar to that evident- and successful- in craft apprenticeship." (Brown et al, 1989; Klabbers, 1990)

[18] Another interesting issue that is being referred to by the phrase 'knowing in action' concerns the question as to whether people change their representation of the situation (or problem) on the basis of activities that they are engaged in (e g. in a participative policy modelling context), or the communication that is taking place For a discussion of this issue, the reader is referred to Klabbers, 1986.

> but by the quantity and quality of domain-specific knowledge." (Gagné, in: Breuer, Tennyson, and Lippert, 1989, p. 30)

Or as Breuer, Tennyson, and Lippert (1989, p. 7) put it:

> "Man kann annehmen, dass die Möglichkeiten für Differenzierungs-, für die Integrations-, und für Kreativitätsleistungen um so wahrscheinlicher werden, je umfassender die Wissensbasis von Individuen ausgebildet ist." (Breuer, Tennyson, Lippert, 1989, p. 7)

As to whether the participants do have the domain knowledge that is required to participate in problem-oriented simulations in order to be able to differentiate and integrate their knowledge (Tennyson, Thurlow, and Breuer, 1988) remains to be seen. However, it is expected that, due to the fact that most of the study's participants have been working in the domain of the Dutch Health Care system for some years, most of them do have the domain-specific knowledge that is necessary to develop and improve their strategic knowledge, that is knowledge of the system dynamics background of participative policy modelling.

## 2.7 SUMMARY

In the present chapter, an overview was given of the policy making process to start with. This to be in the position to understand where and why in the policy making process participative policy modelling can be of any help. Based on the description of the policy making process, it was concluded that participative policy modelling probably seems to have a lot to offer to the first (conceptualization) stages of the policy making process. Firstly, because of the importance of these stages when dealing with ill-structured problems. Secondly, because of the characteristics of the participative policy modelling method itself - participative policy modelling as a cognitively based support tool, focusing on the way in which participants look upon a policy problem.

Following the description of ill-structured problems, a list of features that policy methods dealing with ill-structured problems should have to be able to provide real support, was developed (based largely on Mason and Mitroff (1981)). It was stated that these policy methods should be participative, integrative, adversarial, managerial mind supporting, and context-dependent or situational.

Next, the most important features of system dynamics modelling were discussed. This because participative policy modelling is based on system dynamics modelling and the elements of system dynamics thus can be used to arrive at criteria by means of which the effects of participative policy modelling can be assessed.

With respect to the evaluation of policy methods such as participative policy modelling, it was stated that with a few exceptions, hardly any empirically-based research had been carried out so far. As a consequence, it was decided to use the Vennix study as a frame of reference and compare the present study's approach, method, model, and outcomes to the approach, method, model, and outcomes of his study wherever possible.

Finally, the distinction between domain-specific and strategic knowledge was introduced to be able to distinguish between content-dependent and content-independent elements of one's conceptualization of a policy problem. This to be in the position to examine which aspects of one's conceptualization policy makers change in order to be able to deal with ill-structured complexities, and to determine which of these aspects is supported mostly by the participative policy modelling method.

In the next chapter the present study's theoretical model will be focused upon, that is, an overview will be given of the research questions that will be answered in the present study and the criteria that are used for this.

# CHAPTER 3: THEORETICAL MODEL

## 3.1 INTRODUCTION

Having explained that the present study will be concerned with the evaluation of participative policy modelling in terms of *system dynamics* criteria, account needs to be taken of what exactly these criteria are and how they relate to the research questions concerning the individual and inter-individual change, to arrive at a clear understanding of what is being evaluated in the present study. A description of the 'how', that is, the method by means of which this evaluation is to be carried out, will be given in Chapter Four.

As far as the present study's objectives are concerned, a distinction must be made between the evaluation of the effects brought about by the participative policy modelling and the process by means of which this is being accomplished. The present study not only aims to evaluate the product, that is, the way in which the participants look upon the problem, but also aims to examine the process by means of which these products are brought about.

Once this distinction between product- and process-evaluation has been presented, an outline will be given of the criteria, derived from the world of system dynamics, that will be used to assess the impact participative policy modelling is having on the way in which the participants look upon the problem; to assess the effects produced by participative policy modelling. Note that to classify these effects, use will be made of the distinction between domain-specific changes (that is, changes related to the content of the conceptualization), and domain-independent or strategic changes (that is, changes related to the system dynamics format or structure of the representation).

Following the description of the product-evaluation, account will be taken of two additional research questions, that do not directly fit into the basic research model of individual and inter-individual change. The first additional question concerns the relationship between the present study and the study carried out by Vennix. The second one is concerned with the awareness participants have of the way in which other participants look upon the problem that is focused upon in the sessions. It will be included because of the difference that exists between shared understanding and knowing that shared understanding has been established.

Finally, a description will be given of the potentially confounding variables that will be included in the present study. The chapter is summarized by a theoretical model in which the experimental condition (that is, the participative policy modelling), research questions, and potentially confounding variables are presented in relationship to each other.

## 3.2 RESEARCH QUESTIONS

Two types of questions can be distinguished. The first type, concerning the evaluation of the <u>effects</u> (i.e. summative evaluation) of participative policy modelling, has already been described in some detail. However, in order to arrive at a theoretical model that can be used to empirically assess these effects, a more detailed description of these effects is required. The second type of research question, focusing on the <u>process</u> (or method) rather than the effects (i.e. formative evaluation), has not yet been accounted for in the present study. Information about how the process is perceived by the participants, that is, the perception of the various elements and sessions that make up the participative policy modelling method, is valuable in that it may assist in the explanation of the results. Information on the perceived usefulness of the various components may, for example, be used to differentiate between people who have and people who have not improved their conceptualization of the problem. As such, the process-variables may be used as specifying variables. Moreover, by taking into account the participants' evaluation of the process (e.g. usefulness, time-investment, and difficulty of (parts of) the program), improvements can be made to enhance the method's impact on the participants and its contribution to the policy making process.

### 3.2.1 EVALUATING THE EFFECTS OF PARTICIPATIVE POLICY MODELLING

As described in the foregoing chapter, the present study aims to assess the impact of participative policy modelling on the way in which a problem is being *conceptualized* (that is, represented, looked upon) by the participants. Consequently, the present study is concerned with the conceptual rather than the instrumental impact of the participative policy making method (cf. figure 1.1 of the introduction).

As far as the differences in conceptualization (i.e. differences in mental maps) are concerned, a distinction can be made between *individual* changes and *inter-individual* changes. Individual changes concern the changes individuals as individuals make in the way in which they look upon the problem due to participation, whereas inter-individual changes focus on the way in which the individual mental maps are related to each other, i.e. what these maps have in common in their representation of the problem. It is important to make this distinction because, as explained in Chapter 2, policy making is often carried out by more than one individual, as a result of which differences in perspective are likely to exist (in particular when participants come from different parts of the organization). Arriving at a shared understanding (overlap in the way in which the problem is conceptualized) is an important step forward in the process of deciding on what to do. Because changes in the individual mental maps not automatically result in an increase in the amount of overlap among the participants' conceptualization (participants may change in different directions), individual and inter-individual change need to be distinguished.

Regarding individual change, it is expected that participants will enrich their conceptualization by taking part in the participative policy modelling sessions. This enrichment is not defined in terms of criteria that stem from the world of 'policy theory', but in terms of criteria that stem from the very own characteristics of the participative policy modelling method itself, that is, the underlying 'system dynamics' tradition. Stated differently, it is expected that individual participants will enrich their conceptualization of the problem at hand, by incorporating elements of the system dynamics tradition, and thus describing the problem in a system dynamics way. Although a brief description of the features of system dynamics has already been given in Chapter Two, a more elaborate description is required to arrive at the criteria on the basis of which we can assess empirically whether or not participative policy modelling is capable of changing the way in which people think about, define and structure a problem.

The inter-individual change participative policy modelling aims to bring about, concerns the establishment of a so called 'shared understanding'. It is expected that having the participants work together, discuss and share their ideas in the process of building a generally agreed upon model (conceptual and/or computer model) of the problem at hand, they do not only acquire a better knowledge and understanding of the way in which other participants look upon the problem (get to know their perspectives), but also integrate the perspectives that are bound to be somewhat different.

The two major research questions, concerning the evaluation of the effects brought about by participative policy modelling, can be summarized as follows:

1. *Does participative policy modelling change (for the better, i.e. 'enrich') the way in which individual participants look upon a problem?*
2. *Is participative policy modelling able to bring about an increase in the elements participants have in common in the way in which they look upon the problem?*

In order to arrive at criteria that can be used to assess the degree to which the participants have changed their (domain-specific and strategic[19]) representation of the problem as individuals and/or as a group, a more detailed description of the system dynamics tradition will be given. The first step in this description will be a brief overview of the background of the system dynamics approach. Once the basic ideas of system dynamics have been outlined, a more detailed description of the elements that are essential for a system dynamics point of view will be presented. This to arrive at a list of features that can be used to assess the degree to which our subjects have been able to acquire a system dynamics representation of the problem at hand. However, because use will be made of the distinction between domain-specific and strategic changes in conceptualization, the description of the criteria will be followed by a model relating the criteria to these two forms of knowledge, to arrive at a clear understanding of the kind of change, in terms of which criteria, is being looked for in the present study.

## *System dynamics thinking*

As stated, system dynamics was developed at MIT during the 50s, primarily by Jay W. Forrester. He brought together ideas from three fields that were relatively new at that time: control engineering (the concepts of feedback and self-regulation), cybernetics (the nature of information and its role in control systems), and organizational theory (the structure of human organizations and the mechanisms of human decision-making) to study the dynamic behaviour of what can be called complex, nonlinear multiloops feedback systems (Meadows and Robinson, 1985). To examine and reduce the complexity of these nonlinear multiloops feedback systems, it was felt that one should not break up the problem into more manageable but relatively independent subproblems (Oud, 1978), because the interaction between the system components is more important than the components themselves[20] (Forrester, 1961). The behaviour of complex systems is to a large extent the result of the relationships between the elements rather than the elements themselves. As a consequence, system dynamicist are very reluctant to split the system into subsystems for it may well be that the relationships between the subsystems are of vital importance for the resulting dynamic behaviour of the main system (Vennix, Smits, Geurts, 1989). In other words, in examining complex systems, one should take into account both the components (subsystems) that constitute the total system and the system itself as a whole, realizing that the whole is more important than the sum of its parts (Thissen, 1984). It is this holistic approach to complex problems[21] that makes system dynamics particularly suited to examining the dynamic behaviour of complex problems in the context of policy making and policy development (Klabbers, 1976; Oud, 1978). The system dynamics's ability to reduce the problem's complexity without reducing it into (mono-disciplinary) subproblems (Keuning, 1973; Oud, 1978) is based on the fact that it relies heavily on concepts such as 'system', 'information feedback control', 'elements', 'relationships', and 'boundary', that are of a meta-disciplinary rather than a domain-specific nature. As such, system dynamics can be regarded as a (methodological)

19 As long as the relationship between the system dynamics criteria for assessing a change in conceptualization and the distinction between domain-specific and strategic knowledge has not been explained in more detail, the effects of participative policy modelling will be referred to in terms of 'changes in conceptualization' or 'changes or enrichment of the way in which people look upon the problem' rather than in terms of a change in domain-specific and/or strategic conceptualization It is only when the relationship between the system dynamics criteria and these two aspects of conceptualization has been elaborated upon, that the two research questions described above, will be formulated in terms the distinction between domain-specific and strategic change

20 From a system dynamics point of view, it is the use of the notions of 'system' and 'boundary' that enables one to reduce the complexity of a problem in a non-reductionistic manner For a more detailed explanation, the reader is referred to the sections that follow

21 The relationship between system and problem will be explained in one of the sections to come

instrument (Keuning, 1973) or a way of thinking about or looking at the world in an attempt to understand the world's complexity (Checkland, 1981, Klabbers, 1975) in which concepts of any discipline or field of thought can be included (Meadows, 1980).

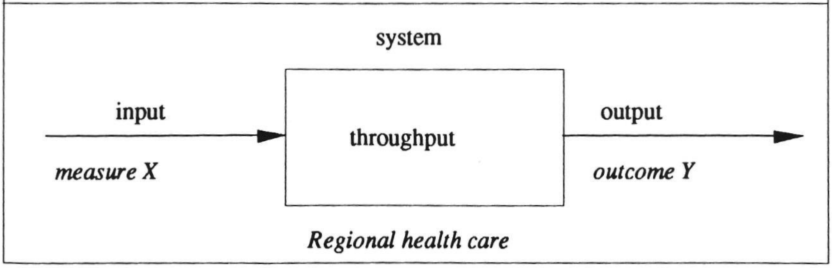
Having explained that system dynamics in general can be considered as a multi-disciplinary, holistic approach to arrive at a shared and more in-depth understanding of complex problems, and to facilitate the communication about those problems (Keuning, 1973; Oud, 1978), account can be taken of the key elements or concepts that make up this particular approach and that help reduce the complexity of ill-structured policy problems. How system dynamics aims to bring about such a reduction in complexity will be discussed in the paragraphs that follow.

One of the most important concepts of the system dynamics approach, as reflected in its name, is the concepts of 'system'. *System* dynamics aims to understand the dynamic behaviour of a part of reality by modelling it as a *system* (Oud, 1978). Although many different definitions of systems have been given[22], most of them stress the idea of 'entirety' by using words like 'entity', 'set', 'whole' and 'black box' and the notion of interrelationship using words such as 'interaction', 'interdependent', and 'relationship' (Keuning, 1973). The use of these notions to define a system is illustrated by the definition provided by for instance Kaufmann (1980), according to which a system can be defined as *"a collection of elements which interact with each other to function as a whole"*.

The advantage of the use of such a system's concept, is that is makes people realize that within the system many elements are related to each other and that it is the combined effect of these inter-related variables that is bringing about the (problematic) behaviour of the system. Moreover, thinking about the problem in terms of a system helps to limit the number of factors and relationships that are taken into account to explain and tackle the problem, for one primarily concentrates on factors that belong to the system. The notion of 'system' thus leads to the concept of 'boundary' as a means to decide whether or not to include particular elements in the representation of the system (Kramer, 1978; Rutges, 1976). As such, the boundary serves to distinguish between the key elements of the representation that form the internal network of endogenous variables and the relatively less important exogenous variables. Or as Forrester (1969, p. 17) puts it: *"anything that is not essential to creating the behaviour of interest is, by that lack of essentiality, on the outside in the unspecified environment"*.

The most simple way to represent the idea of a system as a whole, separated of its environment by a boundary is to depict it as a (black) box on the basis of which input is transformed into output:

Figure 3.1: Representation of a system




The elements and relationships of the system that are required to specify how conditions at the input will be transferred to the output, are commonly called the transfer function (Forrester, 1961). An example of such a transfer function, is the set of elements and

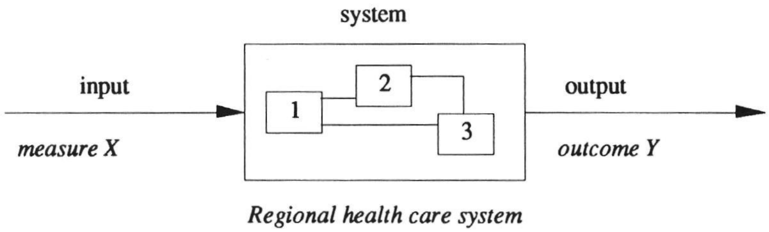
22 A literature research carried out by Keuning (1973), resulted in thirty different definitions of a system.

relationships representing the regional Dutch health care system by means of which the results (output) of a particular measure to reduce the costs of health care (input) can be examined. If the relationship between input and output is not a simple mathematical one, as is the case in our example of the relationship between measure and costs of health care, the transfer function is called a model. It is because of this definition and use of the concept 'model' that the construction of a model is considered as a central element of the system dynamic way of studying the dynamic relationships between in- and output, that is, the examination of the dynamic behaviour of the mediating system.

Having said that complexity of complex (social) systems (or complex problems, for present-day applications stress the importance of the problem in deciding which elements to include in and which elements to exclude from the model; a problem rather than a system should be modelled) can be reduced by representing the problem in terms of a system, thereby using the notion of a boundary to include or exclude elements to arrive at a representation of the model that is capable of explaining the dynamic behaviour of the problem at hand, account should be taken of the fact that because of the counter-intuitive behaviour of many of today's complex systems, explanation and understanding can only be acquired when a closer look is taken at the content of the system under consideration. The black box which can be held responsible for transferring the system's input into output, needs to be opened to arrive at an understanding of which elements and relationships (the so-called descriptive domain-specific content of the model) are in which structure (the so-called normative system dynamics format) are required to bridge the gap between in- and output.

One of the first steps in opening the black box to acquire a deeper understanding of the transformation processes that take place between in- and output, consists of dividing the total system into subsystems[23], in an attempt to structure the large number of elements and relationships that constitute the transfer function or model of the system. To illustrate the effect of box-cutting, we have applied it to the present study's model of the Dutch health care system leading to a system of the Dutch health care in which three subsystems have been distinguished:

Figure 3.2: Representation of a system




1 general practitioner subsystem
2 medical specialist subsystem
3 hospital subsystem

Note that not only the relationships within the subsystems are to be taken into account, but also the relationships between the various subsystems, for they also play an important role in the production of the behaviour of the subsystems and the total system. It is because of this awareness, that system dynamics can be considered as holistic in nature[24].

23 This process is called box-cutting or reticulation (Oud, 1978).

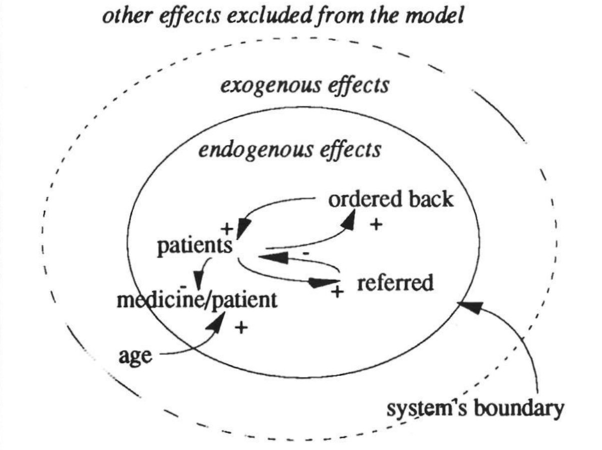
24 Another reason why system dynamics can be considered as holistic in nature is because of its ability to include aspects from any discipline or field of thought.

So one of the first things that can be seen by opening the system, that is, by looking carefully at the model that is being used to represent the system at hand, is its domain-specific *content*, that is, the elements and relationships that are being used to describe the problem or system. From a system dynamics point of view, three major aspects of this content can be distinguished: endogeneity, exogeneity, and multi-disciplinarity.

Endogeneity of content refers to those elements (concepts and relationships) of the description of the system that fall within the boundary of the (sub)system under consideration as a result of which they both affect and are affected by the other elements. It is this endogenous aspect of the content of a system that system dynamicists consider to be most important for the generation of the behaviour of the system (Forrester, 1961, p. 52; Meadows, 1980, p. 34;)

Exogeneity, by contrast, concerns the influence that elements that are themselves not affected by the (sub)system have on the (sub)system's behaviour (Forrester, 1969, p. 18). Exogenous variables help to understand how the internal system of a model reacts to a particular kind of behaviour pattern in the environment external to the system (Forrester, 1961, p. 112). This external environment can be another subsystem, or the environment of the system as a whole, depending on the level of abstraction that the content is being looked upon. To illustrate the difference between endogenous and exogenous variables, reference will be made to the Dutch health care system, the system that is focused upon in the present study. The number of patients treated by the medical specialist, for example, can be regarded as an endogenous variable, for it is affected by concepts such as 'number of patients referred by the general practitioner' and 'number of patients referred back from the hospital'. The concept 'ageing population' however, is considered as an exogenous variable, for although it does affect the Dutch health care system (older patient for example, require a more intensive treatment), it is assumed that within the context of our problem, no causal relationship between the Dutch health care system and the average age of the population does exist. The difference between endogenous and exogenous variables in relationship to the notion of system's boundary can be depicted as follows:

Figure 3.3: Endogenous and exogenous factors




The third important aspect of the content of a system (or model of the system) is the degree to which multi-disciplinary concepts or relationships are included. Multi-disciplinarity is, as already indicated, one of the two ways in which the holistic nature of the system dynamics tradition can be expressed[25]. It contributes to the holistic nature of a representation by virtue of its willingness to include concepts from any field of thought, as a result of which the problem is not necessarily reduced into one particular discipline; system dynamics as a non-reductionistic problem-driven rather than discipline-driven approach to complex, ill-structured problems: *"Since the objective is to include those factors that influence the answers sought, the basis of model building cannot be limited to any one narrow classification of intellectual discipline"* (Forrester, 1961, p. 61). One of the advantages of such a broad conception of the problem is, as already mentioned in Chapter 2, that the likelihood of jumping at conclusions can be reduced.

Having described the content of the system in terms of the elements and kind of elements that are required to represent the system or problem that is focused upon in a non-reductionistic way, account must be taken of the aspects of the representation that serve to understand and explain the dynamic behaviour of the system, that is, their behaviour over time. Based on the description of the content of the system, assumptions must be made about its "structure", meaning the general nature of the interrelationships within the model as representation of the system (Forrester, 1961, p. 58), to arrive at an understanding of the dynamic behaviour of the system. Note that the factors on the basis of which the dynamic behaviour of a system can be explained, are not related (although they cannot exist without a descriptive domain-specific knowledge or content) to a particular application or domain. They are domain-independent and concern the system dynamics format or organization of the domains of application.

The factors that, from a system dynamics point of view, determine the characteristics of information-feedback systems, that is, their behaviour over time, are structure, amplification, and delays (Forrester 1961, p. 130, p. 348).

As far as the structure is concerned, one of the most important aspects of the structure of a system (or model of the system), is the feedbackloop: *"systems can be seen as feedbackprocesses having a specific and orderly structure* (Forrester, 1969, p. 12). Or as Meadows (1980, p. 34) puts it:

> "The primary assumption of the system dynamics paradigm is that the persistent dynamic tendencies of any complex social system arise from its internal causal structure (...). The central concept that system dynamicists use to understand system structure is the idea of two-way causation or feedback."

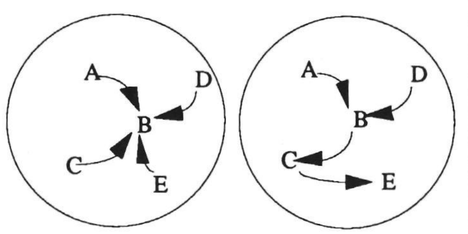
In order to understand, explain, and predict the behaviour of a model (or system), system dynamicists look for feedbackloops (recursive causality) in the representation of the problem. Once they have found a feedbackloop, they determine whether the feedbackloop at hand is a positive or a negative one. Positive feedbackloops amplify deviations or disturbances around the loop. They are characterized as destabilizing, growth producing or self-reinforcing. The dynamic behaviour that positive loops may bring about is exponential in nature (exponential growth or decay). Negative feedbackloops by contrast, do not reinforce but weaken the changes in behaviour, they attempt to negate deviation from some equilibrium or goal state. As a consequence, they are often called goal-seeking or stabilizing loops.

Another aspect of the structure that needs to be taken into account concerns the non-recursive causality of the model. Non-recursive causality is important in that the average length of the chains of singular cause-effect relationships can be considered as an indication of the dynamic potential of the conceptualization for it is only because of these chains of causal relationships that dynamic behaviour of the model can be brought about. Generally speaking, an increase in the average length of non-recursive chains can

25 Recall that it is the breadth of the conceptualization, indicated by the degree to which reference is being made to subsystems and/or elements that do not fall within the boundary of the total system, the so-called exogenous variables and relationships, which serves as the second way in which the holistic point of view of a conceptualization of the system or problem, can be expressed.

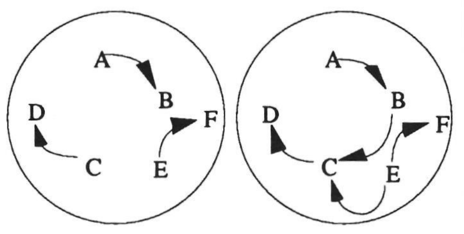
be seen as an increase in the awareness that changes in one element of the system do not only result in a change in one other element, but that this next element is bringing about a change in other elements as well. An example of a chained reasoning is: 'if A increases, this will bring about a decrease in B, as a result of which C and D decrease as well, which, in turn will lead to an increase in E'. The difference between chaining and non-chaining can be expressed as follows:

Figure 3.4: Chaining




poor conceptualization rich conceptualization

Figure 3.5: Connectedness




poor conceptualization rich conceptualization

The last aspect that, from a system dynamics point of view, needs to be included in the consideration of the structure of a model, is its connectivity. Connectivity refers to the degree to which the concepts of the conceptualization are related to each other. It is important in that it serves to measure the degree to which the concepts of the model are actually being 'used' in the model[26]. The usage of a concept is based on the frequency with which the concepts in the model are being employed as part of a causal relationship. If the ratio of concepts to relationships is low (i.e. a low connectivity), a relatively high number of concepts is used to describe a relatively low number of causal relationships, as a result of which the concepts have low causal explanatory power; they contribute only little to the understanding of the dynamics of the model. Because system dynamics modelling aims to reduce the system's complexity by including only those elements in the model that are important for the explanation of the problem, a relatively high degree of connectivity is to be preferred, for it means that less elements are needed to describe the problem.

As shown in figure 3.5, the number of relationships thus plays an important role in the assessment of the 'relatedness' of the representation. Note that because a low number of meaningful elements is to be preferred to a large number of meaningless elements in order to arrive at a concise (and thus better) system dynamics representation of the problem, a conflict with the previously described notion of 'content of a system' seems at hand. The potential conflict shows that just expanding one's knowledge base (or domain-specific knowledge) will result in a richer domain-specific conceptualization. However, in order to improve on the dimension of strategic knowledge, that is, knowledge in action of system dynamics, additional (format) requirements such as the structure of the knowledge base, need to be met. So rather than increasing one's score on the 'content' dimension by simply adding various elements and relationships that are only loosely connected to the already existing representation, or improving one's connectivity score by just removing elements, resulting in a very concise but poor (in terms of content) representation of the problem, one should try to enrich one's conceptualization by improving the content of the conceptualization (a broader, more in-depth, or multi-

---

26 Note that the criterion 'connectedness' is not exclusively used by system dynamicists to assess the quality of the conceptualization. Prawat (1989, pp. 6-7), for example, states that *"one thing that has become clear from the expert-novice research is that the expert's knowledge base is organized around a more central set of understandings than novices"*. And *"some ideas are more meaningful than others - which is to say, they allow for a richer set of connections."*

disciplinary view on the problem) while maintaining the connectivity of the model's structure at at least the same level. The result of this improvement would be a rich but still highly connected 'theoretical model' of the problem.

How the various criteria are related to each other (Is there a conflict between content and format related criteria? Is it possible that people improve on one aspect of their system dynamics representation of the problem at the cost of another aspect, so that the quality of their profile in total does not change?), will be discussed in more detail in the chapter concerning the outcomes of the present study. In that chapter, the discussion will take place in light of the distinction between domain-specific and strategic knowledge.

Having considered the aspects of the structure of the model that are important to take into account when examining the model's behaviour, attention can be focused on amplification as another important aspect of the system's dynamic behaviour. *Amplification is manifest when an action is more forceful than might at first seem to be implied by the information inputs to the governing decisions* (Forrester, 1961, p. 15-16). *It is a response from some part of the system which is greater than would at first seem to be justified by the causes of that response. Amplifications arise in the policies that define the decisions that control the rates of flow* (Forrester, 1961, p. 62). An example of such a policy that is affecting the flow of patients in the Dutch health care system, is the general practitioner's policy to regulate his or her workload by, among others, adjusting the percentage of people that are referred back for another consult. Depending on the way in which factors such as structure, delay, and possible information-distortion (e.g. by averaging figures) interact, amplification may arise from this particular policy. For example when, due to an information-delay, general practitioners decide to increase the number of patients that are referred back at a time that the workload is high rather than low (or vice versa), amplification will be found. To understand the dynamic behaviour of the model, an awareness of where, how, and why amplification is taking place is required.

The third and final aspect of the model that can be considered a crucial determinant of the dynamic behaviour of the system, is the so called time delay. The concept of 'delay' is closely related to the notion of 'amplification', for it is, among others, the presence or absence of a delay[27], that determines whether or not amplification is taking place in a particular chain of causality. Increasing the amount of time that is required for a material or information flow to go through a system, to go from one element to another (i.e. increasing the delay), reduces the effect of that particular chain on the overall behaviour of the system. Note that delays specify the relationship between two elements rather than the elements themselves. An example of a delay is the time that is required for a general practitioner to realize that a change in workload is structural rather than due to incidents such as a flu epidemy. Phrases such as 'an increase in X will result in an increase in Y *in the short run*' and '*after four weeks* the increase of P will bring about a decrease in Q' are often found in descriptions of a problem to express the persons awareness of a particular delay between two variables that have a causal relationship.

Summarizing the above description of the system dynamics approach to complex problems, a distinction is made between the content and the format of the conceptualization (or model) of the system. The content of the system concerns the description of the system that is being focused upon in terms of the elements that need to be included. The way in which the content is structured, that is, the format of the conceptualization, is taken into account to explain the dynamic features of that particular content.

Having described the elements on the basis of which the change in conceptualization will be evaluated, account must be taken of the direction of the change that participative policy modelling is expected to bring about. For instance, is it expected to result in an increase or a decrease in the average length of the causal chains that are used to describe

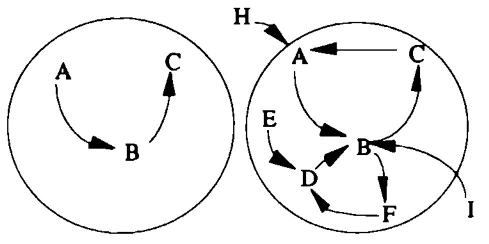
---

27 For a description of the way in which delays or variability of delays can create amplification, the reader is referred to Forrester, 1961, p. 348.

the problem? For each of the eight elements discussed above, this particular question will be answered, and a brief explanation will be given of the reason(s) why.

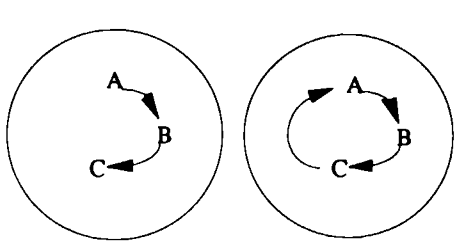
In general, as far as the first two aspects of the way in which the content of the system is represented are concerned, it is expected that participation will lead to an increase in both the number of concepts and number of relationships that are used to represent the system, for the system that we are dealing with is a complex one[28]. Note that the relationship between an increase in concepts and relationships used to describe the problem and an increase in the knowledge one has of the issue at hand is not characteristic for the system dynamics way of thinking only. Many other approaches believe that an increase in content shows that the person at hand has acquired a more refined understanding of the problem. Since people can improve both their endogenous and exogenous knowledge independently (an increase in knowledge about endogenous variables does not automatically lead to more or less knowledge about exogenous variables or vice versa), these two criteria will be included separately. To illustrate the difference between good and limited understanding of the problem in terms of endogenous or exogenous concepts and relationships:

Figure 3.6: Number of endogenous and exogenous concepts and relationships




poor conceptualization rich conceptualization

Figure 3.7: Feedbackloops




poor conceptualization rich conceptualization

The third aspect of the content that is important to assess how much progress participants have made in their conceptualization of the problem concerns the multi-disciplinarity of their perspective on the problem at hand. Improving the content of the conceptualization also means that elements from different disciplines are taken into account. By doing so, it contributes to the holistic, multi-disciplinary character of the system dynamics approach as described in both this chapter and Chapter Two.

Regarding the format of the conceptualization, it has been made clear that feedbackloops are important in explaining and understanding the behaviour of the system. Consequently, the more feedbackloops are included in the conceptualization of the problem, the better the conceptualization is from a system dynamics point of view[29]. The difference between a conceptualization with and without such a feedbackloop is depicted in figure 3.7.

Regarding the criterion of chaining, it has already been said that an increase in the average length of the non-recursive chains of causal relationships should be strived for because they indicate an increase in awareness that changes in one element can bring

28 Obviously, the more knowledge one has to start with, the more difficult it will be to add concepts and relationships to the already existing conceptualization in a meaningful way. However, since the problem that is being dealt with is a complex, ill-structured one, as a result of which no existing model or theory can be used to describe the problem, it is likely that an increase in the number of concepts and relationships can be considered as an increase in knowledge rather than a non-meaningful extension of the conceptualization.

29 Obviously, more does not always mean better. It holds within a particular range for the purpose of a system dynamics approach to complexity is to reduce the complexity by limiting oneself to the most important relationships (and/or feedbackloops).

about changes in elements that are not directly related to each other, that is, in a direct causal relationship. It is expected that participation will lead to an increase of this awareness, and thus will bring about an increase in the average length of these chains.

It has already been indicated that the connectivity of the representation should be improved. As will be explained in more detail later, the ratio of concepts to relationships will play an important role in this.

The last two elements, amplification and delay, however, will be combined in the evaluation of the effects of participative policy modelling. The reason for this is twofold. First of all, as to whether people are aware of amplification is difficult to assess in the description of a problem, for participants would have to state explicitly that the response from some part of the system is greater than would at first seem to be justified by the causes of that response (Forrester, 1961), to be able to conclude that amplification indeed has been incorporated in the particpants' conceptualization. Hence it was decided, because factors such as feedbackloops and delays play an important role in creating amplification (Forrester, 1961, p. 348), to use the element of 'delay' as an indicator for potential awareness of the so-called time-phase relationships in which amplification and delays work together to determine the dynamic behaviour of the system[30]. An increase in the number of delays thus can be considered as an increase in the system dynamics character of the description of the problem.

Based on the characteristics of the system dynamics approach, a number of criteria have been presented by means of which an assessment can be made of the degree to which participants have changed the way in which they conceptualize the problem at hand. In this conceptualization (in terms of system dynamics: model, or transfer function), content and format elements can be distinguished. The former is concerned with the domain-specific knowledge that is included in the description of the problem, that is the (number and kind of) concepts and elements incorporated in the description. The latter, however, concerns the way in which these domain-specific elements are structured in order to arrive at an understanding of the problem's complexity. As such, it focuses on the way in which this domain-independent knowledge can be organized and used in the context of policy making. As to whether participants have been able to adopt a system dynamics approach to deal with the problems that they were confronted with, will be assessed on the basis of the criteria that are related to the way in which the domain-specific knowledge has been organized. The distinction between content and format criteria thus mirrors the previously discussed distinction between domain-specific knowledge and strategic knowledge.

As far as the individual change is concerned, that is, the 'enrichment' of the participants' individual conceptualization of the problem, the combination of the system dynamics criteria (content and format) and domain-specific versus strategic knowledge, thus results in the following list and classification of criteria on the basis of which the effects of participative policy modelling will be assessed:

30 For a more detailed description of the way in which phase shifts between two variables can be related to both the delay and smoothing functions that contain amplification, the reader is referred to Forrester 1961, Appendix I, pp. 421-425.

Figure 3.8: System dynamics criteria

| Domain-specific knowledge | | |
|---|---|---|
| content | exogenous | concepts relationships |
| | endogenous | concepts relationships |
| | multidisciplinary | concepts relationships |

| Strategic knowledge | | |
|---|---|---|
| format | structure | feedbackloops |
| | | chaining |
| | | connectivity |
| | amplification | time-phase relationships |
| | delays | |

To assess the amount of change that is taking place at the ***individual*** level, posttest scores are compared to pretest scores, and participants are asked to complete a questionnaire. How this evaluation will be carried out, will be explained in more detail in Chapter Four.

Regarding the ***inter-individual*** change, it has already been said that the present study aims to assess the increase in the elements participants have *in common* in the way they think about the problem. What these elements are is something that needs to be considered in more detail. A first answer to the question of what the elements are on the basis of which it is determined whether or not the participants increasingly look alike in their representation of the problem as a result of taking part in participative policy modelling, is that it will concern both the content and the system dynamics format of the representation, that is, the domain-specific and strategic knowledge that is included in the conceptualization. As a consequence, the question of whether participation is able to bring about an increase in what the participants have in common while conceptualizing the problem, consists of two separate questions:

1. *Is participative policy modelling capable of bringing about an increase in what the participants have in common as far as the system dynamics format or strategic aspects of their representation is concerned?*

2. *Is participative policy modelling capable of bringing about an increase in what the participants have in common as far as the domain-specific content of their representation is concerned?*

The first question can be answered using the system dynamics criteria developed earlier on. If the variance of the degree to which elements such as feedbackloops, chains, delays and quantifications are included in the conceptualization is decreased, the participants' mental maps will be more alike, as far as their system dynamics format is concerned. Formulated in this way, a collective avoidance of the use of feedbackloops could be considered as an increase of 'communality' or 'shared understanding'; the understanding that feedbackloops are not worthwhile taking into account. However, since we believe that 'inter-individual change' should not take place at the cost of 'individual' change, i.e. the individual level of system dynamics thinking should not be decreased to arrive at a generally agreed upon representation of the problem, there is an additional requirement that inter-individual system dynamics change has to meet: the individual posttest scores on system dynamics thinking should not be lower than their pretest scores. Preferably, they should be higher than the pretest scores so that the participants both would have acquired a shared understanding and an understanding of what system dynamics thinking is like. In concreto, if participative policy modelling is capable of bringing about an individual change in the system dynamics way of thinking about the problem without decreasing the average level of system dynamics thinking (the average level may remain the same if some individuals improve while others deteriorate), and the variance on the system dynamics criteria is reduced, then one can conclude that people look more alike in the way they conceptualize the problem; a shared understanding is arrived at, as far as the system dynamics format of the representation is concerned.

Regarding the second question, the question concerning the content of the representation, a shared understanding is being arrived at when people use the same concepts and relationships to describe the problem at hand. However, it is important to note that this equivalency has to be an *equivalency of meaning* rather than of terminology. The reason for this is that we do not want to create an artificial similarity which would occur if, say, participants decide to stop using the word 'PC' and use the word 'computer' instead. If this would happen, it would seem as if the participants had acquired a shared understanding while in reality they had already been referring to the same part of (conceptual) reality[31]. In coding the pre- and posttests, account will have to be taken of this potentially confounding influence of equivalency of terminology.

As a consequence, to assess the change in content of the individual and inter-individual conceptualization, the number of concepts and relationships will play an important role. However, to evaluate the inter-individual change, one also has to take into account the kind of concepts and relationships that are being used to avoid the above mentioned artificial unification.

Concerning the content of the conceptualization, Vennix (1990) suggests that one additional research question should be taken into consideration, the question of how many of the concepts and relationships that belong to the conceptual and computer model created during the sessions are indeed incorporated by the individual participants. In the present study, this question will be addressed at both the individual and inter-individual level. As a consequence, we will examine the amount of concepts and relationships of the external model (conceptual or computer model) that are included in the individual conceptualizations <u>and</u> discuss whether or not they use the same concepts and relationships to conceptualize the problem at hand. The above mentioned research questions can be summarized as follows:

Figure 3.9: Research questions

<table>
<tr><th></th><th>strategic knowledge</th><th>domain-specific knowledge</th></tr>
<tr><td rowspan="2">individual change</td><td rowspan="2">increase of average score on criteria (structure, amplification and delays)?<br>(I)</td><td>increase of average score on criteria (exogenous, endogenous and multi-disciplinary)? (IIa)</td></tr>
<tr><td>incorporation of concepts and relationships of the external model (exogenous, endogenous and multi-disciplinary)? (IIb)</td></tr>
<tr><td rowspan="2">inter-individual change</td><td rowspan="2">increase of average score on criteria and decrease of variance on criteria (structure, amplification and delays)?<br>(III)</td><td>reduction of variance of used concepts and relationships (exogenous, endogenous and multi-disciplinary)? (IVa)</td></tr>
<tr><td>reduction of variance of incorporated elements of the external model (exogenous, endogenous and multi-disciplinary)? (IVb)</td></tr>
</table>

On the individual level, the present study aims to assess as to whether the participants increase both their domain-specific and strategic knowledge in relationship to the problem at hand. Note that regarding the domain-specific knowledge, we would like to know whether people improve on the exogeneity, endogeneity and multi-disciplinarity dimensions, and assess how much of the external model, created and used by the group

[31] As to whether they referred to the same part of conceptual reality (connotation) or empirical reality (denotation) is a very interesting philosophical issue, but falls beyond the scope of the present study.

of participants during the participative policy modelling sessions, is incorporated in the individual representations. Concerning the strategic knowledge, however, attention is being focused upon the degree to which the participants have managed to apply the principles of system dynamics to their own conceptualizations, that is, organize the content of their representation according to the characteristics of system dynamics thinking.

On the inter-individual level, three questions will be asked. First, regarding the content, how much the participants have increased their commonality of thinking in terms of the aspects of domain-specific knowledge discerned above and the content of the external model. Regarding the strategic knowledge, the present study aims to determine whether their strategic approach is becoming more alike in that they all approach the problem from a system dynamics point of view. For this, the very same aspects distinguished at the level of individual changes will be used.

### 3.2.2 EVALUATING THE PROCESS OF PARTICIPATIVE POLICY MODELLING

Having focused on the evaluation of the effects brought about by participative policy modelling, account needs to be taken of the evaluation of the process, the so-called formative evaluation. As mentioned before, examination of the way in which participants look upon the process (that is, the method itself) can be used to improve the method (if necessary) and to assess the effects in more detail. The ways in which the method is experienced by the participants can be used as potentially confounding variables, that is, as variables that may disguise the relationship that exists between treatment and effect(s).

To meet these two objectives, it was decided to develop a questionnaire, divided into three parts, reflecting the three major components of the participative policy modelling program: preparatory workbook, small-group activities, and plenary presentation and discussion. Participants were asked to indicate how interesting, easy, and useful each of the three components was, whether they felt they had learned something from it, what their opinion on the duration of these components was. Based on these scores, a total score for each of the components, and a total score on each of the four aspects (summing the scores over all three components) can be arrived at as follows:

Figure 3.10: Evaluating the program (formative evaluation)

| | workbook | small-group activity | plenary session | |
|---|---|---|---|---|
| duration | | | | total = ........... |
| interesting | | | | total = ........... |
| useful | | | | total = .......... |
| easy | | | | total = .......... |
| learned from it | | | | total = .......... |
| sum of the five aspects | | | | |

Moreover, participants were asked to indicate how much they felt the workbook was related to their job, and whether or not the workbook had met its goal of preparing the participants for the session to come. For a more extensive description of the questionnaire, the reader is referred to Appendices 1 to 4.

### 3.2.3 ADDITIONAL QUESTIONS

In addition to the research questions concerning the change in content and format of the individual and inter-individual conceptualizations of the problem, two more issues are addressed in the present study.

The first issue concerns the comparison of the present study to the Vennix (1990) study. Despite the differences between the present study and the study carried out by Vennix (cf. Chapter Two), comparisons between the two studies will be made whenever it is felt appropriate. Not only as far as the findings based on the pre- and posttest measures are concerned, but also regarding the results that stem from the evaluation of the process. As such, the Vennix study will serve as a kind of reference since it is one of the very few studies that can be compared to ours with respect to its empirically based evaluation.

The second question focuses upon the awareness participants have of the way in which the other participants look upon the problem. Awareness is different from homogeneity in that one can be aware of the way in which other people think without adapting to their frame of reference and vice versa. Awareness need not include adaptation, however, it seems logical to assume that adaptation does require some kind of awareness. It is difficult to believe that without some kind of awareness of the other participants' point of view, adaptation can take place. It is even harder to believe that without such an awareness people can state that they have adapted in such a way that their level of homogeneity has increased. In other words, questions number 33 and 34 (Appendix 3), referring to the amount of knowledge people have of the other participants' point of view can be viewed as a prerequisite for question number 35 (Appendix 3), in which the participants are asked whether they have changed their minds in such a way that the level of homogeneity has been raised.

To sum up, the research questions that will be focused upon in the present study are divided into questions concerning the products of participative policy modelling, and questions related to the process itself. Regarding the product-related questions, five questions will be distinguished. Four of them can be classified using the dimensions of content vs system dynamics and individual vs inter-individual change. The fifth question considers the awareness of the other participants' point of view, which should, logically speaking, precede inter-individual changes. The process-evaluation concerns issues such as duration, easiness, level of interest, and usefulness. These process variables will, together with background variables such as age, gender, educational background, be used as potentially confounding variables in the present study's theoretical model. These potentially confounding variables will be used to arrive at more specific knowledge about the relationship between the present study's experimental condition (participative policy modelling), and its dependent variables (the product-related research questions).

## 3.3 POTENTIALLY CONFOUNDING VARIABLES

Now that the research questions of the present study have been outlined, account needs to be taken of the variables that may affect the results the participative policy modelling has on the individual and inter-individual conceptualization of the problem, the potentially confounding variables mentioned above. Following Vennix (1990), these potentially confounding variables will be divide into three categories: background characteristics, evaluation of the process, and time investment.

Although we do not have explicit hypotheses regarding each of these three kinds of variables, including them in our analyses can provide us with additional information regarding the effects brought about by the participative policy modelling. It may tell us, for example, that specific groups (e.g. those who are more experienced, or those who did appreciate the participative policy more) change their conceptualization in particular, or that certain groups (e.g. those who are less experienced, or less positive about the program) do not change at all.

Regarding the first category of variables, the background variables, it was decided to include variables such as age, gender, department, education, number of years working for the company, area of expertise, organizational background (pre-merger organization), and percentage of working time spend on policy making. In line with Vennix (1990, pp. 75-76), intelligence, cognitive complexity, and heuristic competence were not included. Moreover, cognitive learning style was also excluded because the validity and reliability of the learning style questionnaire (Kolb, 1976; 1981) has been questioned (Basuray, 1982; Huber, 1983; Vennix, 1990).

The reason why educational background was included as a potentially confounding variable, is because the participants of the present study fall into two categories as far as their educational background is concerned, those who studied medicine and those who did not[32].

Department serves as a potentially specifying variable because the participants stem from two different departments; a medical department and a financial-legal department[33]. Note that the above-mentioned two educational categories match the two departments in that the people with the medical background all worked at the medical department while those with a non-medical background worked for the financial-legal department.

The area of expertise was included to examine how much knowledge the participants had about areas that did not belong to their own area of expertise to start with (that is, measured on the pretest). Moreover, based on information regarding the area of expertise, assessment could be made as to whether the participants improved in their own area of expertise, in other areas than their own field of expertise, or in both.

The second kind of potentially confounding variable concerns the amount of time spent on the program. Using the number of hours spent on preparation and the amount of time spent in the program itself, a measure of the total amount of time invested can be obtained. Based on this, a distinction can be made between the intention-to-treat and an on-treatment group (cf. Chapter 4).

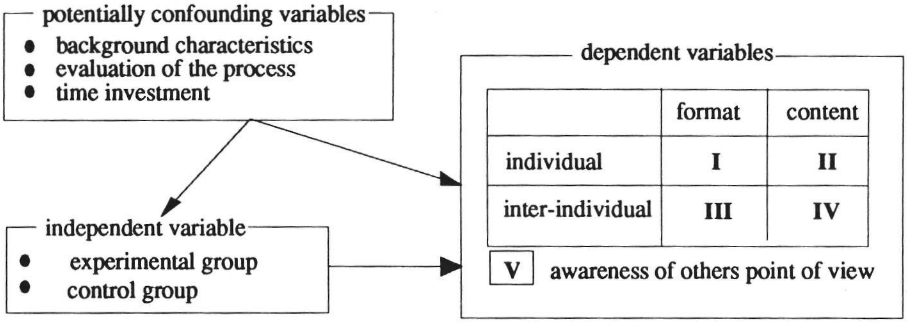
The last group of potentially confounding variables stems from the process evaluation carried out by the subjects themselves. Using scores on variables such as 'easiness', 'usefulness', or 'appreciation of the subgroup activities', a distinction can be made between say, people who did and people who did not appreciate the subgroup activities. Again, it should be mentioned that we do not have any specific hypotheses regarding the effects of these variables, but we have included them to elaborate upon the potential causal relationship between participative policy modelling and conceptualization of a (policy) problem.

To summarize the present study's theoretical model can be expressed as follows:

32 For a more detailed description of the participants, the reader is referred to Chapter 4.

33 The questions regarding the department and the pre-merger organization were not included in the questionnaire because we were able to answer these questions ourselves.

Figure 3.11: Theoretical model




## 3.4 SUMMARY

In the present chapter, an overview has been given of the major elements included in the present study. To start with, the basic research questions were discussed, and a distinction was made between research questions that concern the outcomes or effects of participative policy modelling, and research questions that are concerned with the process of participative policy modelling itself.

Using the system dynamics frame of reference, indicators for the product evaluation were arrived at. Following the distinction between domain-specific and strategic knowledge, a theoretical model was constructed in which individual indicators were related to either the category of domain-specific knowledge or the category of strategic knowledge. Moreover, a description was given of the background variables, the variable time-investment, and the process variables. This to not only get an idea of the amount of time invested in the program, the kind of people taking part in the program, and the way in which they program was evaluated by the participants, but also to be in the position to assess the potential confounding effects these kinds of variables may have on the relationship between the independent and dependent variables. Finally, an overview was given of the theoretical model, depicting the way in which the major elements are related to each other.

In the next chapter, the way in which the research questions outlined in the present chapter will be answered will be focused upon in detail.

# CHAPTER 4: DESIGN, METHOD AND OPERATIONALIZATION OF THE VARIABLES

## 4.1 INTRODUCTION

Having described the research questions that will be answered in the present study, account must be taken of the way in which this will be accomplished. To be in the position to evaluate the products brought about by the participative policy modelling approach and the process by means of which this is accomplished, a research design has to be selected that meets the requirements of both the world of research methodology (so that proper research can be done), and the world of the participating organization. The tension between these two requirements will be illustrated by the distinction between intended and actual design. Prior to the description of these two designs however, some methodological implications of the situational or historical status of knowledge for the design and interpretation of the present study will be discussed. Following the description of the intended and actual design, the methods used to collect the data will be focused upon - an overview will be given of the pre- and posttest, and the questionnaire by means of which both the products and process of participative policy modelling will be evaluated. Next, a presentation will be given of the way in which the present study's dependent variables and potentially specifying variables will be operationalized, to arrive at a list describing what variables are being measured in what way. Finally, a supplementary description will be given of the stimulus or treatment, that is, the participative policy modelling method. However, since the activities of the policy method have already been described in Chapter Two, a description will focus on the program's time-table to arrive at an understanding of who is doing what and when.

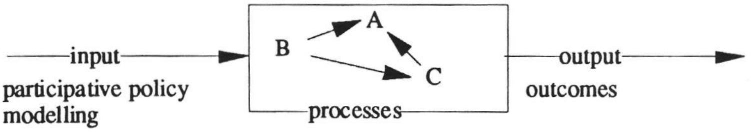
## 4.2 METHODOLOGICAL IMPLICATIONS OF THE NOTIONS OF 'HISTORICISM' AND 'SITUATIONALISM'

Recall that in Chapter Two a distinction was made between a rationalistic and a historicistic approach to the production and use of knowledge in the policy making process. According to the rationalistic approach, the knowledge that is being arrived at is context-independent and transcends the historical situationality. As such, it aims to find rules or laws that hold irrespective of time and location. The historicistic or context-dependent approach to knowledge, by contrast, considers knowledge to be dependent of the situation in which it is being arrived at; knowledge that holds in one particular situation, not necessarily holds in another situation.

As far as the model-building process is concerned, the distinction between these two epistemological approaches reflects the distinction between traditional system dynamics and its participative or interactive version. In the former, the complexity that is inherent to many of today's policy problems, is being dealt with by having expert model-builders construct a model which is capable of unequivocally predicting the output of a system on the basis of the system's input. The knowledge that is being produced thus can be considered as relatively independent of the person(s) who constructed the model; the expert model-builders are regarded as neutral observers rather than participants. The participative or interactive model-building, by contrast, incorporates the historical context of the situation in the process and product of knowledge production by having policy makers interact with the system, that is, adjust and analyze the system, as a result of which they will begin to realize that the system's output (behaviour of the system) is the result of both the model's initial conditions and the adjustments they themselves have made to the model. The relationship between in-and output (that is, the knowledge of the system being arrived at) thus varies depending on the people who contribute to the model-building process. Moreover, since the people who participate in the process of building the model are elements of the system themselves (they are actors in the field of health care), conceptualizing the health care system in a particular way will affect the way in which they behave in that system as well, and thus will modify the very same system. As a consequence, both the traditional distinction between subject and object of knowledge, and the distinction between virtual and real system, that is, between the model created by the participants, and the real-life system they are part of, disappears[34] (Klabbers, 1988). The knowledge of the system that is being arrived at, is knowledge from an insider's perspective, taking into account the system's internal state in explaining the relationship between in- and output.

Having explained that the model-building process has moved from a relatively rationalistic to a more historicistic approach to knowledge production, some methodological implications for the way in which the present study is carried out must be discussed for believing in the historical context of knowledge and knowledge production in the context of model-building, obviously should be reflected in the approach to bring about knowledge about the participative policy modelling process. The study itself, can be considered as the attempt to find a (or the?) transfer function between input and output, that is, the relationship between the independent and dependent variables, the participative policy modelling method and the changes in the conceptualization, it is bringing about:

Figure 4.1: The present study from a system's point of view




To assess the effects of participative policy modelling, that is, to determine as to whether a causal relationship between in- and output exists, both a rationalistic and a historicistic approach can be adopted.

Within the rationalistic approach, context-independent knowledge of the relationship between the participative policy modelling method and the participants conceptualization is being focused upon. Datamodels are often used (Klabbers, 1988) to arrive at a deterministic explanatory model and to assess the quality of the knowledge thus being arrived at, use is made of the notion of internal validity.

The historicistic approach, by contrast, aims to take into account the circular processes that may exist between the in- and output, thereby accounting for the system's history in the explanation of the relationship between participative policy modelling and problem conceptualization. By having an eye for the processes by means of which input is being transformed into output, rather than concentrating on the correlation between in- and output only, the historicistic approach aims to open the black box in between in- and output to acquire a kind of insider's perspective to account for the relationship between in- and output. The historicistic approach thus seems to be concerned with the external

34 Note that the primary reason why the distinction between virtual world (shared and constructed model) and real or actual world (external system) is disappearing, is not because the virtual one is a valid representation of the actual reality. The reason why the distinction between the two is vanishing, is because of the self-referential character of the reflection that is taking place in the construction of the model. In constructing a model, participants reproduce their own reality, that is, (re)structure the reality in which they operate.

validity of the knowledge being arrived at, rather than being preoccupied with the internal validity of the study's conclusions.

As far as the present study is concerned, it should be mentioned that the discussion as to whether a rationalistic or a more historicistic approach should be taken to examine the effects brought about by participative policy modelling, has not been our primary concern. It is our impression though, that this study can be considered as falling within the rationalistic approach for it primarily aims to detect a relationship between in- and output. It is only on the basis of the presence or absence of such a relationship that intervening processes (for example by means of the specifying variables) will be focused upon. This is not to say, however, that we are not aware of the fact that the historical setting in which the present study has been carried out, may have had an impact on its outcomes, and it is because of this awareness that an attempt will be made in the chapter concerning the interpretation of these outcomes, to account for the differences in conceptualization in terms of some of the processes that took place at the time of the treatment at the participating organization and its direct environment. However, in spite of the attention that will be paid to these potentially intervening processes, we still believe that the present study should be considered as remaining within the rationalistic tradition, as illustrated by the weight that is attributed to internal validity in the selection of the most appropriate research design - it is only on the process of carrying out the research that we have moved somewhat from a purely rationalistic approach to a historicistic-rationalistic one.

## 4.3 INTENDED VERSUS ACTUAL DESIGN

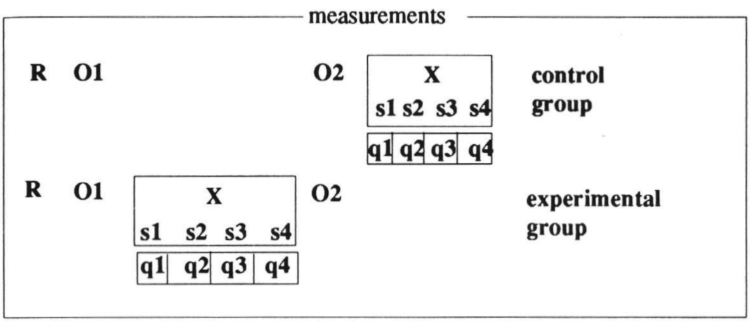
In the description of the design that will be used to assess both the effects brought about by participative policy modelling and the process itself, a distinction will be made between intended and actual design. The intended design reflects our conception of the way in which the present study should (and could) be carried out, based on both methodological and practical (organizational) considerations. However, as will be explained in more detail, adaptations to this intended design had to be made on request of the organization[35], resulting in what can be called, the actual design, i.e. the design that eventually has been used to assess the effects and process of participative policy modelling

[35] Note that it is only because of the inclusion of processes like this, that a more profound understanding of the application of participative policy modelling can be acquired. Without the description of the interaction between the researcher and organization taking place prior to and during the participative policy modelling sessions, a limited perspective on the method and its effects would be obtained.

## 4.3.1 INTENDED DESIGN

In order to assess the <u>effects</u> of participative policy modelling, it was decided to use the following design:

Figure 4.2: Intended design




s1 to s4: session 1 to session 4
q1 to q4: questionnaire 1 to questionnaire 4

in which: O1=pretest
X=participative policy modelling
O2=posttest
R=randomization

The design is based on the 'untreated control group design with pretest and posttest (Campbell & Stanley, 1966; Cook & Campbell, 1979)[36]. However, it differs from it in that the control group is also offered the very same treatment (X), though at a later point in time. This to enable all members of the organization taking part in the program to experiences the same policy making support.

One of the most important reasons for selecting this particular design is that, because of its randomly assignment of participants to either the experimental or the control group, it is able to take care of many of the threats to internal validity (Cook & Campbell, 1979, p. 56). As a consequence, the effects of potentially disturbing variables can be ruled out without having to specify all of them in advance. However, since the number of participants that can be assigned randomly to one of the two groups is relatively small (thirty-two), it is recommended to check as to whether the two groups are indeed equivalent with respect to some of the variables that can be considered as potentially disturbing. Another advantage of this particular design is that since the treatment is also offered to the control group, threats such as demoralization of the non-treatment group, imitation of the treatment, and compensatory equalization and rivalry can be ruled out.

In spite of these advantages, however, there is one threat to internal validity that is not automatically controlled for by the 'untreated control group with pretest and posttest design': the threat of differential mortality. Differential mortality, that is differences between the two groups due to differences regarding the number or kind of people that drop out of the program still remains as a potential threat to the study's conclusions. However, as suggested by Cook and Campbell, examination of the proportion of people that actually took the posttest and analysis of the pretest scores of those who completed

36 The control group is offered the very same treatment posterior to the posttest since the participating organization insisted that both departments would participate to the participative policy modelling program. In order not to affect the basic controlgroup design, it was decided to provide the control group with a treatment once they had completed their posttest.

the posttest, will give an indication of whether the dropouts differed across groups on the background variables that are most likely to affect the posttest scores.

The degree to which the results that stem from this design can be generalized to particular target persons, settings, and times, that is, the degree to which the result based on this design will have an external validity, is questionable (Campbell & Stanley, 1966, p. 8). However, since our primary goal is to reveal the causal effects produced by participative policy modelling ***applied in a field setting***, rather than the examination of whether or not this causal relationship can be generalized to a more general population, internal validity is considered to be a more important threat to our study than external validity is.

Two more aspects of the intended design need to be dealt with: the number of organizations participating to the study, and the way in which the groups have been put together.

Regarding the number of organizations taking part in the participative policy modelling program, we originally had in mind to have more than one organization participate in the project to increase the study's external validity. However, it soon became clear that because the program's lead time (including pretest and posttest measures) was equal to six months, and a preliminary computer model had to be built before the sessions could be carried out, the number of participating organizations had to be limited to one. Obviously, this affects the degree to which the study's findings can be generalized to other settings and populations. However, as said above, the study's primary goal is not to generalize to other setting and populations, but to assess the degree to which a causal relationship between participative policy modelling as independent variable, and individual and inter-individual change as dependent variables, can be established in a real-life situation with real policy makers rather than students.

Whether or not to assign individual participants to homogeneous or heterogenous groups is important from a research methodological point of view[37]. However, one should also take into consideration the research questions themselves and determine what kind of groups are required to study these questions best. Since the present study is aimed at determining the degree to which participants that have a somewhat different background, increase the commonality in the way in which they look upon the problem due to their participation, both the experimental and control group should consist of people that differ in background (e.g. education, department, and experience). Consequently, it was decided to randomly assign the participants to the two groups, and to check as to whether the two departments (medical and non-medical department), and the two pre-merger organizations[38] were equally represented in each of the two groups. This because it was felt that organizational background and departmental background could be important variables in that they can affect the way in which people look upon a particular (policy) problem.

Summarizing, a pretest posttest control group design seems to be the most appropriate design to assess the effects brought about by participative policy modelling, because of its ability to rule out most threats to internal validity. Moreover, the design is in accordance with the requirements that stem from the research questions : it allows for differences in background (departmental and thus educational, and organizational background). It also meets the organization's most important requirement that no one should be excluded from the program; all potential participants should take part in the program.

---

37 Heterogenous groups that are as much as possible equivalent to each other are to be preferred over homogenous groups that are very much different to each other. This to reduce the error due to differences between persons (Cook & Campbell, 1979, p. 47).

38 The participating organization had been involved in a merger process. Since it was felt that both departmental and pre-merger organizational background could account for differences in perspective, all participants were matched on these two variables, prior to being assigned randomly to the experimental or control group.

Note that the effects of participative policy modelling are not only measured by means of the pre-and posttest measures. Participants are also asked to rate how much they have learned and changed. This is done by having participants fill out a questionnaire at the end of each session[39]. The reason for this is that by asking the participants whether participation has resulted in a change (individual and inter-individual), comparisons between the outcomes of the pre- and posttest and the results from the questionnaire can be made to see as to whether any differences between the two exist.

To evaluate the process responsible for bringing about the effects, that is, the participative policy modelling itself, the very same questionnaire will be used. Participants will be asked to indicate whether they have found the session useful, long, interesting, and difficult.

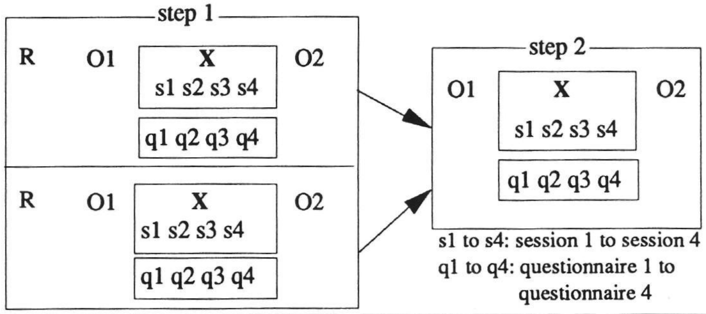
Despite of the fact that we expected the intended design to be able to meet the needs of the organization and the requirements of scientific research at the same time, last minute adaptations to the design had to be made. Why the design had to be changed, and what these changes were, will be described in the next section concerning the actual design, that is, the design that has actually been used in the present study.

### 4.3.2 ACTUAL DESIGN

The intended design could not be employed in the present study because, rather than dividing the participants into two groups and having one group participate in the program at a time, the participating organization insisted that all participants should take part in the program at more or less the same time. It was felt that the participants had to partake at the same time to increase the likelihood of a true shared understanding, i.e. an understanding of the activity (participative policy modelling) and the problem focused upon, by most of, if not all, the people working at the two participating departments.

Note that this is not to say that everybody had to participate in one and the same group, or exactly at the same time. The organization did not accept a postponed treatment for a control group, but was willing to accept two groups participating in the same activities (sessions), say, two days apart from each other[40]. The reason why the organization was willing to accept this alternative was first of all because we convinced them that real participation is more difficult to establish in one large group than it is in two smaller groups. Moreover, the organization had become aware that if everybody would participate at exactly the same time, staffing problems would occur at both departments. To avoid this, it was decided to change the intended design into the following actual design:

Figure 4.3: Actual design




39 The questionnaire will be described in more detail in section 4.4, where the methods used to evaluate the process and product of participative policy modelling are focused upon.
40 For an overview of the program and corresponding time-table, see section 4.6.

In figure 4.3, it is shown in step 1 that the difference between the experimental and the control group, based on the time at which people take part in the program (X), will disappear if all participants take part in the program at the same time. The procedure that was used to assign the participants to either the control or the experimental group can still be employed to divide the total number of participants into two groups that are as much as possible identical to each other (in particular on the variables 'departmental background' and 'pre-merger organization') and internally heterogenous with respect to the same two variables. As a consequence, the 'R' standing for randomization, still holds for the actual design, because the participants were assigned randomly to one of the two groups after we had matched them on the variables 'departmental background' and 'pre-merger organization'. It is important to maintain the randomization procedure to reduce the likelihood of differences between the two groups for differences between the groups may result in different interaction patterns between the facilitator and participant and among the participants themselves, as a result of which the two groups cannot be aggregated into one single group.

Since the differences between the two groups are limited (if there are any significant differences at all), and we do not intend to compare the two groups, or look upon them as two separated evaluation studies, it was decided to consider them as one large group (this would increase the number of participants in the experiment dramatically) despite the fact that three of the four sessions were carried out separately[41]. The resulting design is called the One-Group Pretest Posttest Design (Campbell & Stanley, 1966, p. 8, Cook & Campbell, 1979, p 99), and is depicted in step 2 of figure 4.3

Now that the main reasons for changing from the intended design to the actual design have been given and the nature of the changes have been presented, attention must be focused upon the consequences these changes (may) have on the research questions to be answered by the present study. Having changed the design, we have to consider whether the research questions have to be changed as well. In order to answer this question, account must be given of the potential threats of the design first, for it is only on the basis of these threats that we can determine whether or not the actual design is able to meet the present study's research objectives. For each of these threats, we will discuss whether or not the threat is implausible in the present study's setting or can be convincingly ruled out via direct measurement, thus resulting in an interpretable design.

One of the first threats that may affect the results of the One-Group Pretest Posttest Design is *history*, that is, events other than the treatment (X) may affect the study's outcomes

Regarding the present study, history can be ruled out as far as the strategic knowledge or system dynamics way of looking at the problem (both individually and inter-individually) is concerned. The reason for this is that we believe that a true system dynamics modelling perspective can only be acquired by taking part in a kind of system dynamics seminar (or participative policy modelling program), which certainly did not exist elsewhere in the Netherlands at the time of our treatment.

41 The reason why it is believed that the two groups were subjected to the same treatment despite the fact that they participated in two separate groups for most of the time, is first of all because the groups are as much as possible identical as a result of which more or less identical experiential and educational background is drawn upon in the process of building the model Secondly, as will be explained in the description of the timetable, both groups were provided with the same preliminary model as a consequence of which their discussion of what to include or exclude in the final model hardly differed The adaptations that were suggested during the sessions were almost identical for both groups Not only because both groups were almost identical, and they both received a more or less identical stimulus (the preliminary model), but also, and this is the third reason why the processes in the two groups can be considered as one and the same process, because the participants intensively discussed the activities they had been participating in at lunch-breaks, and the revisions proposed by one group were used as an input for the discussions of the other group as well, as a result of which the computer model used in the third session (based on the revisions proposed in both groups) was completely identical for both groups.

As far as the domain-specific content of the representation of the problem is concerned, note that the influence of exogenous factors (extra-curricular activities, that is, activities that do not belong to the treatment such as reading the newspaper or watching the news) can never be ruled out completely. From a theoretical point of view, it is not possible to determine as to whether the effect is brought about by the treatment alone, or that the effects must be accounted for by some external events as well. However, in the present study, we have no reason to believe that such a systematic exogenous influence has indeed affected the outcomes and we thus assume that if an increase in, say, the number of concepts and relationships of the conceptualization is found, it has to be attributed to the model built as part of the participative policy modelling program rather than to external influences (history).

The second threat to internal validity that needs to be dealt with when using the One-Group Pretest Posttest Design, is *statistical regression*. Since the actual design does not exclude any potential participants from participation (say, because they had only little experience, or low scores on the pretest), the most common form of a regression artifact for this design can be ruled out (Cook & Campbell, 1979, p. 100).

Another potential threat related to this design is the threat of *maturation*. As will be made visible in the time-table outlined in section 4.6, the lead time of the project is about six months, and it may well be that an increase in knowledge is related to an increase in maturation (say, experience) rather than to the treatment itself. Although we do not believe as to whether six months is enough to produce a maturation effect in a matter as complex as the reduction of the costs of health care, an attempt will be made to rule out the effect of maturation by calculating the correlation between experience (years with the firm) and scores on the pretest (Cook and Campbell, 1979, p. 101).

*Testing* may be another threat in that participants can be made aware of what needs to be learned and thus do better on the posttest. However, in the present study, the participants were asked to fill out a pretest almost three weeks before the first session, suggesting that the pretest was not related to the treatment at all. Moreover, the participants were told explicitly that the test did not ask for a right or wrong answer but simply was designed to capture their individual point of view regarding some topical issues in the health care system. Hence, since the participants were not told what the present study's research objectives were, it is expected that the effect of the pretest on the posttest can be neglected.

The threat of *instrumentation* can be ruled out since the measuring instrument or scorers used in the present study were not changed. Moreover, the absence of a control group rules out additional threats such as *demoralization of non-treatment group* or *compensating rivalry*.

The above discussion of the potential threats to both the intended and actual design can be summarized as follows:

Figure 4.4: Threats to internal validity

| + = *ruled out as a problem (implausible)*<br>+/- = *for some elements a potential threat*<br>? = *needs to be determined by analysis* | | maturation | history | testing | instrumentation | regression | compensation |
|---|---|---|---|---|---|---|---|
| intended design | R O1 O2 X<br>R O1 X O2 | + | + | + | + | + | + |
| actual design | O1 X O2 | ? | +/- | + | + | + | + |

Figure 4.4 shows that for most of the potential threats, it is believed that the case has been made that they are implausible in the present context. However, note that although we strongly believe that most of the threats that may apply to the actual design have been ruled out successfully, causal inferences that stem from the actual design are not as robust as those arrived at on the basis of the intended design.

Having discussed the effects that may be brought about by the change in design, we are now in the position to answer the question of whether or not the research questions need to be adjusted as well. Since the present study's primary goal is to examine if and to what degree there exists a causal relationship between treatment (i.e. participative policy modelling) and effect (change in conceptualization), rather than being able to generalize the study's findings to other settings and populations, the research questions as described in Chapter Three, can still be answered even now that the actual design is being used rather than the intended one. Changing a control-group design into a design without any control, does not imply that changes can no longer be examined (provided that we have a pre- and a posttest). However, it does mean that in interpreting the results, one has to take into account the limitations to validity as described above. In other words, we do not have to change the research questions now that we had to give up our intended design and use the actual design instead, as long as we take into consideration the effects this change in design has on the robustness of the causal inferences that can be drawn.

To summarize, the present study had to change from an intended to an actual design for reasons provided by the organization. Although this may affect the internal validity of the outcomes (it is our believe that most of the threats can be made implausible), a change in the research questions is not required. Consequently, the present study can be characterized as a case study whose objective it is to both <u>apply</u> the participative policy modelling method to a real-life situation and to carry out (some) <u>research</u> on real-life participants. The first objective serves to increase our understanding of the method's strength and weaknesses when applied in reality for future development of the participative policy modelling method. Knowledge about the process (say, attainability, degree of interest, usefulness and so on) will be used for this. The second objective refers to the research carried out to evaluate the effects brought about by the participative policy modelling method. It enables us to compare some of the results of the Vennix laboratory study (1990) with the present study's field research.

Now that the most important research questions have been described (Chapter Three), and an overview has been given of the design actually used to answer these questions, account must be taken of evaluation methods that will be used to assess both the process of participative policy modelling and the effects that are being brought about by it. Hence the next section will provide a description of the questionnaire and pre- and posttest that will be used to assess the process and products of the participative policy modelling program.

## 4.4 METHOD

In this section a description is given of the methods that will be used to assess how the participants look upon the program, whether they have made any changes in their conceptualization of the problem at hand, and the degree to which they have acquired some understanding of how the other participants look upon the problem at hand. In line with Vennix (1990), a questionnaire and a pre-and posttest will be used to assess the effects and process of participative policy modelling. However, in contrast to that study, no use will be made of a multiple-choice test to measure the amount of 'universally accepted' knowledge the participants have about the subject at hand. The reason for this is, as already described in Chapter One, because no theoretical framework to classify knowledge claims as correct or wrong exists with respect to ill-structured problems.

### 4.4.1 THE QUESTIONNAIRE

In order to examine the participants' perception of the participative policy modelling program, it was decided to have the participants complete a questionnaire at the end of each session. The main aspects included in this evaluation concern the level of interest of the program, its usefulness, the program's difficulty, whether something has been learned from it, and whether the session has taken too much time (for an overview of the questionnaire, the reader is referred to Appendices 1 to 4). In addition, people are asked to express their feelings with respect to the preparatory text they had been reading, and the degree to which the components of the program were related to each other in a meaningful way. The questionnaire was designed in such a way that by summing the scores within each of the components (i.e. preparatory text, small-group activities, plenary presentation and discussion), the program itself could be evaluated. However, by summing the scores over the components within each of the five aspects, an idea of how interesting, useful, difficult, long, and learningful the program is, can be arrived at as well. The first summation thus will results in a total score for each of the three components, whereas the second one will lead to a total score on each of the five aspects mentioned previously. Note that a five-point scale is used, as a result of which scores higher than three are considered to be positive, expressing one's appreciation of the program in terms of say, usefulness or interestingness. Scores lower than three, by contrast, indicate a person's dislike of the program in terms of the very same qualifiers, and finally, a score of exactly three is considered as a neutral one, expressing that the person is neither positive nor negative about the qualification at hand. The questions were put in the form of a semantic differential, to limit the number of items to be included: rather than asking first whether the workbook had been easy <u>and</u> then asking them whether the workbook had been difficult as well (saying that the workbook had not been easy is not the same as saying that it had been difficult), it was decided to include both extremes into one item, thereby reducing the number of items substantially.

The questionnaire was not only designed to evaluate the way in which the participants felt about the process in terms of both components and aspects, but also to evaluate the effects brought about by that process. The reason why it was decided to ask participants about their perception of the effects, is because it enables us to examine <u>if</u> and to <u>what</u> degree the participants' perception of the effects of participative policy modelling differs from the effects measured by means of the pre- and posttest scores.

Since the questionnaire will be administered prior to the posttest measure, it is important that the items are formulated carefully, in order to avoid sensitization of the participants; the participants should not be made aware of the expected outcomes of the study. To avoid such a sensitization, the items had to be formulated in a relatively general way, and some items were removed from the questionnaire as well (to avoid the questionnaire from becoming too long).

### 4.4.2 PRE- AND POSTTEST

The present study not only aims to examine the *perceived* effects of the treatment, that is, the effects as they are rated by the participants themselves, but also aims to assess the effects in a more objective, that is, participant-independent, way. Hence it was decided to have the participants complete a pre- and posttest.

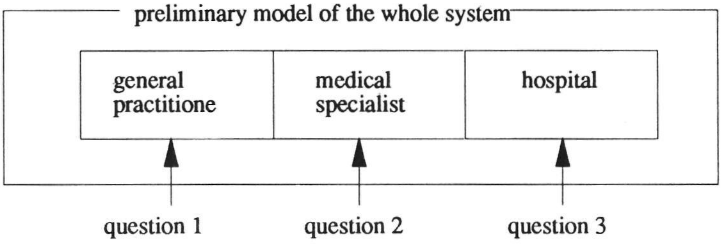
Since a pre- posttest design is also used by Vennix (1990) to assess differences in conceptualization, and comparisons with the Vennix study are an important objective of the present study, it was decided to follow his approach as much as possible.

Besides, Vennix had his participants write a policy note. They were asked to complete one open- ended question concerning the issue at hand. The very same question was administered prior to and posterior to the participative policy modelling sessions. The two policy notes resulting from this pre- and posttest, were subsequently (re)coded using

Axelrod's (1976) procedure[42] to arrive at a data-matrix format so that further analyses could be carried out.

The present study, however, differs from Vennix regarding the theoretical framework that is used to assess the differences in conceptualization. In contrast to the 'policy-theory' frame of reference used by Vennix, the present study employs a 'system dynamics' frame of reference to assess whether any changes in conceptualization have taken place. Because the preliminary model[43], used to start the model-building process and increase its pace, appears to consist of three sub-systems, reflecting three possible areas of expertise, it was decided to include three questions in the pre- and posttest (each related to one particular subsystem) rather than one general question concerning the health care system in its totality:

Figure 4.5: The Dutch health care system divided in three components




To assess the change in scope of conceptualization, participants were asked to answer one open-ended question about each of the three subsystems. The number of elements included in the answer that do not belong to the subsystem the question is referring to, can be used to assess the degree to which participants include exogenous elements in their conceptualization, and thus change the scope of their representation. In other words, if the question is referring to the first section of the health care system, that is, the subsystem of the general practitioner and all the processes that affect him or her, elements that belong to the medical specialist or hospital sections can be considered as relatively 'exogenous' elements. The more elements from the second and third subsystem are included in the answer to the first question, the better the conceptualization will be in terms of exogeneity (obviously within certain limits; elements that stem from that part of the system the question is referring to, should not be excluded from the answer altogether).

Another advantage of using a pre- and posttest that consists of three questions rather than just one, is that it can be used to examine the relationship between area of expertise and quality of answer since the participants were also asked to state which of the three sections they knew most about. Examining the relationship between the claimed area of expertise and quality of answer allows us to see whether their claimed expertise is reflected in their pre-test scores. We obviously expect experts on general practitioners' issues to be able to provide the richest answer to a question related to the general practitioner section of the system. Moreover, it enables us to see whether participants improve their conceptualization (if they improve at all) in the area where they already have a substantial amount of knowledge, or whether most improvements are being made in the areas where there is more room for improvement, that is, the sections where they

42 A more detailed description of the Axelrod recoding procedure will be given in Chapter Six concerning the construction of the current study's variables.

43 In Chapter Five, an overview will be given of the preliminary model.

claim to are less knowledgeable. Note that this issue touches the issue of the relationship between strategic and domain-specific knowledge, as discussed in Chapter Two.

Having discussed the two methods that will be used to evaluate the process and products of participative policy modelling, an overview can be given of what exactly is being measured in order to answer the major research questions outlined in Chapter Three. As a consequence, the operationalization of the theoretical model as outlined in Chapter Three, will be focused upon in the next section.

## 4.5 OPERATIONALIZATION

In this section, a detailed description is given of the way in which the main elements of the theoretical model, as depicted in figure 3.11, will be measured. The presentation will start with the operationalization of the study's dependent variables representing the five types of changes that participative policy modelling is aiming to bring about. This will be followed by the operationalization of the model's potentially specifying variables[44], to arrive at a clear picture of what is being measured for what research questions, or what research questions will be answered on the basis of what measurements.

### 4.5.1 OPERATIONALIZATION OF THE DEPENDENT VARIABLE(S)

Recall that the objective of the present study is to answer the five research questions depicted in figure 3.9 and 3.11. The first four questions are concerned with individual and inter-individual changes in knowledge, either of a domain-specific or strategic nature. The fifth question, by contrast, focuses on a change in awareness or knowledge of a somewhat different nature. Rather than being concerned with the knowledge people have of the subject of the Dutch Health Care system, the fifth research question is concerned with knowledge participants have of other people, that is, the knowledge they have of the other participants' conceptualization of the very same Dutch Health Care system. The dependent variables thus are:

* domain-specific knowledge
* strategic knowledge
* awareness of the others' points of view

With respect to the domain-specific knowledge, the present study examines whether a change has been brought about in the knowledge participants have about the domain of the Dutch health care system, and as to whether their domain-specific knowledge is becoming more alike as a result of their participation.

The strategic knowledge is used to assess whether participants are able to acquire a system dynamics perspective on the problem, and whether they increase their communality with respect to this system dynamics perspective.

Finally, the awareness of the others' points of view is used to assess whether participants have increased their knowledge of how the other participants look upon the problem as a result of taking part in the participative policy modelling sessions.

How the present study's research questions are operationalized, that is, represented in either the pre- and posttest or questionnaire, will be focused upon next. To start with, a description will be given of the way in which the dependent variables have been

[44] Note that, in contrast to the presentation of the theoretical model in Chapter Three, variables such as background characteristics, time-involvement, and evaluation of the process, are called specifying rather than confounding variables. The reason for this is that in the actual design no use will be made of a control group, as a result of which no longer the 'true' differences, that is adjusted for these so-called confounding variables, between two groups, can be examined. These variables, however, still can serve to specify the relationship between dependent and independent variables, that is, act as specifying variables (cf. Segers & Hagenaars, 1980, p. 30).

incorporated in the questionnaire: a description will be given of the items that are included to assess how the participants feel they do on each of the three dependent variables in relationship to the five research questions. Following this presentation, a description will be given of the way in which the dependent variables have been operationalized on the basis of the pre- and posttest.

### 4.5.1.1 USING THE QUESTIONNAIRE

Originally, it was our intention to develop a questionnaire covering all five research questions so that the evaluation of the effects could be based on both pre- and posttest measures and self-ratings. However, as discussed above, because the formulation of the items had to be relatively vague to avoid making the participants aware of the kind of behaviour the researchers were looking for, some of the research questions had to be excluded from the questionnaire.

***Research Question I: Individual change in strategic knowledge***

Recall that to assess the degree to which individuals change the system dynamics format of their knowledge base, it was decided to use the system dynamics elements 'feedbackloop', 'chaining', 'connectivity', and 'time-phase relationship' (cf. figure 3.8). Of these four elements, it was decided to include the first three as follows (for an overview of the questionnaire the reader is referred to Appendices 1 to 4; note that the items 26, 27, and 29 stem from the third questionnaire, presented in Appendix 3):

1. Feedbackloops or recursive causality (question 29): It is expected that the more participants say they are aware of the importance of feedback mechanisms in explaining and predicting the behaviour of complex systems, the more likely it is that feedbackloops are included in their conceptualization of the problem[45].

2. Chaining or non-recursive causality (question 27): To assess whether people have been able to extend the causality in their reasoning about a problem, that is, have increased the average length of the successive cause-effect relationships, we asked them whether the sessions had helped them to explain why certain developments have occurred (and what effects particular measures have resulted in) for being able to do so requires a certain degree of use of cause-effect relationships.

3. Connectivity (question 26): To assess how participants felt about the integratedness of their conceptualization, we asked them whether or not they had discovered any (new) connections between elements of the health care system.

The criterion of time-phase relationship, as combination of the elements of delay and amplification, has not been included in the questionnaire. The reason for this is that it was decided to interpret the impact of this element on the behaviour of a system in terms of its underlying structure in relationship to the element of delay. In other words, amplification was interpreted as the result of the system's structure (in particular its feedbackloops) together with its delays, for together they determine the loop's execution time and thus affect the influence the loop is having on the behaviour of the total system.

---

[45] Obviously, it is realized that a distinction exists between knowing that feedbackloops are important and actually using and incorporating feedbackloops in the conceptualization of a particular domain-specific problem. In Chapter Six, where the construction of the variables is taking place, it will be decided not to include this particular item in the analyses because it is realized that this item need not say anything about the actual use of feedbackloops at all and thus cannot be used to compare self-ratings to researcher-based evaluations (based on the pre- and posttest).

Since it had already been decided to include feedbackloops in the questionnaire, the question rose as to whether the factor 'time' had to be included as well.

The reason why it was decided not to include the element of time in the questionnaire, was because we did not want the participants to realize that time (delays) were considered an important element in the evaluation of their understanding of the dynamics of complex systems. It was already felt that question 29, referring to the criterion of feedback, was too specific to be included, but could not be avoided because of possibility that people use feedbackloops without realizing it (feedback loops are made of singular causal connections which may 'accidentally' form a feedback loop). Since accidental use of 'time-phase elements' (e.g. by distinguishing between long-term and short-term effects in relationship to other developments over time) can be ruled out, it was decided to exclude the 'time-phase' element from the questionnaire.

***Research question II: Individual change in domain-specific knowledge***

***Question IIa: Change in content***

Recall that to assess the quality of a person's domain-specific knowledge, use is made of the concepts of exogeneity, endogeneity, and multi-disciplinarity. To operationalize these concepts, the following items have been included in the questionnaire:

1. Endogenous and exogenous content (question 32): To assess whether participants feel they have changed the content of their conceptualization, (either in an exogenous or an endogenous manner), they were asked whether or not the method had helped them to broaden their perspective on the problem. Note that participants can broaden their conceptualization by either increasing the 'size' of their representation, that is, change the system's boundary (endogenously broad), or increase the number of exogenous elements included in the conceptualization of the problem (exogenously broad). Since question number 32 cannot distinguish between these two ways of increasing the scope of the representation, broad is taken as referring to both endogenous and exogenous improvement of a participant's conceptualization.

2. Multi-disciplinary (question 28): In the present study, an improvement of the multi-disciplinary character of a conceptualization is considered to take place if people incorporate elements in their conceptualization that come from disciplines different from the one in which they were educated. Since the two participating departments differ substantially with respect to their educational background, it was decided to ask the participants whether participation had led to knowledge of the health care system not directly related to their daily work, assuming that their daily work is related to the discipline in which they were educated. For example, people from the medical department, having been to medical school, improve on the multi-disciplinary criterion if they take into account, say, aspects that concern the financial side of the problem. Note that a change in multi-disciplinarity is conceptualized in terms of a qualitative change rather than a quantitative one. A quantitative change in content, i.e. more of the same, is accounted for by question number 32 where a change in content is related to the number of elements that are used in either an endogenous or an exogenous way.

3. Content in general (Questions 30 and 31)[46]: In order to assess to what degree the participants themselves believe the content of their conceptualization has changed,

[46] The questions number 3, 14, 19, 23, and 38 have not been used as indicators of a general change in content for it has not been made explicit whether they refer to strategic or domain-specific knowledge. These items will be used however, to assess how the participants did feel about the program (cf. 4.5.2).

questions number 30 en 31 were included. In question number 30, participants are asked whether or not taking part in the participative policy modelling sessions has helped them to state *more precisely* what is and is not important to the problem of containing the costs of health care. If the method does indeed result in a more precise conceptualization of the problem, participants must have changed the way in which they look at the problem. Note that this improvement can be in terms of multi-disciplinarity, endogeneity, and exogeneity; no particular aspect of the domain-specific content is being referred to. Secondly, participants are being asked whether their interaction with the other participants had provided them with *new ideas and knowledge* (question 31). New ideas and knowledge do not mean a change in the content from one particular point of view (e.g. multi-disciplinarity), but certainly mean that new concepts and relationships have been incorporated in the conceptualizations, indicating that indeed a change has taken place in the domain-specific content of the conceptualizations.

### *Question IIb: Incorporation of the external model*

Research question IIb, concerning the degree to which individual participants incorporate elements of the external model into their own conceptualization of the problem, has not been included in the questionnaire. It was felt that if the participants had been asked as to whether the sessions had resulted in an increase in the number of elements of the external model that were incorporated in the conceptualization of the problem, the impression might have been given that adopting elements from the external model is important and something to strive for, as a result of which the status of the model might have changed from 'free-to-adjust' to 'important-thus-valid'. This would have been in conflict with the participative or interactive policy modelling's objective to arrive at context-dependent or historicistic knowledge, in which participants are given the opportunity to play an important role in the construction of the model and thus are given the opportunity to 'create' their own reality.

### *Research question III: Inter-individual change in strategic knowledge*

Having described in detail the way in which the first two research questions (concerning the *individual* changes in conceptualization), have been taken care of in the questionnaire, account should be taken of the reasons why it was decided not to include the third question, also dealing with the system dynamics format of the conceptualization, but this time on an inter-individual level, in the questionnaire.

The main reason why we decided not to ask the participants whether or not their system dynamics format had become more homogeneous due to the participative policy modelling program, is because of the issue of sensitization discussed before. We believe that there is a difference between asking in a general tone 'has the method succeeded in bringing about (new) relationships between the elements of the health care system and the way in which you look upon the problem?' and asking 'has the method succeeded in bringing about a change in the way in which you look upon the problem so that you and your colleagues have the same (new) relationships in mind when thinking about the problem dealt with in the participative policy modelling program'? Because the latter formulation explicitly refers to a similar kind of change in the way other participants look upon the problem, it is much more obvious that this change is not just one of the many possible ways in which an individual can change, but that it is indeed a preferred way of looking at the problem, showing that we want them to acquire a shared understanding of the problem at hand.

Moreover, it is also doubtful as to whether participants have an idea of the format of the other participants' conceptualization. For most people, knowing the *content* of the others' conceptualization is already a difficult task, let alone that they have an idea of the *format* of the other person's conceptualization. In other words, many people already have difficulty knowing how other people look upon a particular problem, let alone that they know how this conceptualization is being organized or as to whether it represents a system dynamics point of view.

***Research question IV: Inter-individual change in domain-specific knowledge***

***Question IVa: Domain-specific content***

For the very same reason why it was decided not to ask the participants whether their conceptualizations had become more alike with respect to the system dynamics format of their thinking, it was decided to include only one general question regarding the degree to which participants had acquired a domain-specific homogenous representation of the problem. Consequently, it is only question 35 which refers to the degree to which the participative policy modelling method has been able to bring about such a homogeneity or commonality.

## *Question IVb: Incorporation of the external model*

The degree to which participants have become more alike with respect to the external model constructed and used during the model-building sessions has not been included in the questionnaire for the same reasons as it was decided not to include this incorporation of the external model on the individual level. We simply did not want the participants to feel that the external model was more important than their own individual mental model (conceptualization) of the model, because what we are interested in is a change in knowledge in-use (both domain-specific and strategic) rather than a change in knowledge in-theory, that is, knowledge that is not really incorporated and hence is not actively used in the conceptualization of a particular policy problem.

***Research Question V: Awareness of others' points of view***

Notice that the question of whether participants have become aware of the others' points of view will only be answered on the individual level. No attempt will be made to assess as to whether participants have become more alike in the degree to which they are aware of the others' points of views. The reason for this is that participative policy modelling aims to bring about a shared understanding with respect to the domain-specific content and the strategic framework chosen to organize this content, rather than a shared awareness of the other participants' awareness. To measure the degree to which the participants feel they have become aware (on an individual level) of the other participants' points of view with respect to the matter under discussion, questions number 17, 24, 25, 33, 34, 42, and 43 were included in the questionnaire. Questions number 17, 24, 25, 42, and 43 assess this (potential) awareness at each of the individual sessions, whereas questions 33 and 34 by contrast, assess the perceived influence of the first three sessions together have had on the awareness of the others' points of view (e.g. 'have the first three session resulted in...'[47])

To summarize the operationalization of the five research questions ***on the basis of the questionnaire***, figure 4.6 is included to depict the operationalization of the individual changes in conceptualization (research questions I, II, and V). The inter-individual changes are presented in figure 4.7 (research questions III and IV).

47 Recall that there were 4 sessions held in total.

Figure 4.6: Operationalization of the dependent variables using the questionnaire

| *individual level* | | *question and number* |
|---|---|---|
| domain-specific knowledge (II) | general | The participative policy modelling sessions enabled me to state more precisely what is and is not important for the problem of cost containment in the health care system (#30) *(session 3)*<br>By communicating to people with a somewhat different view on the health care system, I have acquired new ideas and insight (# 31) *(session 3)* |
| | endogenous | The participative policy modelling sessions have resulted in a broader perspective on the health care system (#32) *(session 3)* |
| | exogenous | |
| | multi-disciplinary | The participative policy modelling sessions have resulted in an increase in the knowledge I have of aspects of the health care system that fall beyond my daily work-activities (# 28) *(session 3)* |
| | external model | not included |
| strategic knowledge (I) | feedbackloop | The participative policy modelling sessions have made me feel that feedback processes are important in considering a complex system such as the health care system (# 29) *(session 3)* |
| | chaining | Due to the participative policy modelling sessions, I have acquired a deeper understanding of the causes of potential developments and the effects of potential policy measures taken in the health care system (#27) *(session 3)* |
| | connectedness | The participative policy modelling sessions have made me see (different) connections between the various components of the health care system (# 26) *(session 3)* |
| | time-phase relationship | not included |
| awareness of others' points of view (V) | | The assignments and discussions of the small-group session have increased my knowledge of how my colleagues think about the health care system (# 17) *(sessions 1 thru 3)*<br>Altogether, I have acquired more knowledge of the ideas and opinions of the other participants by taking part in this session (# 24) *(session 1 thru 3)*<br>This session has succeeded in making me exchange my thoughts intensively with the other participants (# 25, #43)*(sessions 3 and 4)*<br>Due to the participative policy modelling sessions, I have acquired an idea of how my colleagues think about the health care system (# 33, # 42) *(sessions 3 and 4)*<br>Due to the participative policy modelling sessions, I know about what I agree and disagree with my colleagues (#34) *(session 3)* |

| Figure 4.7: Operationalization of the dependent variables using the questionnaire | |
|---|---|
| *inter-individual level* | *question and number* |
| domain-specific knowledge (IV) | It is my impression that due to the participative policy modelling sessions we have become more alike in the way in which we think about the health care system (#35) (*session 3*) |
| strategic knowledge (III) | not included |

### 4.5.1.2 USING THE PRE- AND POSTTEST

To assess the individual and inter-individual changes in conceptualization, participants not only were asked to fill out a questionnaire at the end of each session, but were also asked to complete a pre- and posttest. Having outlined how the criteria, on the basis of which change will be assessed, are measured in the questionnaire, account will be taken of the way in which the very same criteria are measured by means of the pre- and posttest. Since individual and inter-individual change in conceptualization are based on the very same criteria, no distinction will be made between individual and inter-individual changes in conceptualization as far as the description of the operationalization is concerned. As a consequence, the domain-specific versus strategic knowledge distinction will be used to organize the presentation of the operationalization based on the pre- and posttest rather than the research questions as was the case in the preceding sections.

By asking our participants to complete a pre- and a posttest, that is have them write a policy note prior to and posterior to the sessions, the present study follows the Vennix study. Note however that in contrast to that study, our participants were asked to respond to three policy issues rather than one. This to be able to diversify the scope of the conceptualization (i.e. have them not focus on one aspect of the health care system only), and increase the likelihood that a cognitive map can be extracted from the texts written by the participants. This because it was believed that by having the participants answer three questions, a more substantial description of their knowledge of the health care system would be given. Moreover, as will be explained in more detail later, asking the participants three questions focussing on three different issues rather than one, would enable us to examine as to whether any differences in expertise among the participants existed.

In order to be able to determine whether participants have changed their conceptualization due to taking part in the participative policy modelling sessions, the texts written as part of the pre- and posttest, have to be transformed into analyzable data. For this use will be made of the cognitive mapping approach.

Prior to discussing how the criteria that stem from the system dynamics approach can be measured on the basis of the pre- and posttest, account must be taken of the procedures by means of which the pre- and posttest are transformed into so-called cognitive maps, and the characteristics of these cognitive maps for it is on the basis of these cognitive maps that the operationalization of the current study's variables is taking place. Since the transformation of written statements into a set of causal relationships will be discussed in some detail in Chapter Six, where among others, the validity and reliability of the procedure is focused upon, it suffices to examine the outcomes of those recoding procedures at this stage, for it is on the basis of this outcomes, the cognitive maps, that changes in the conceptualization are being assessed.

Cognitive maps are graph(ical) representations of verbal or written statements in which concepts are being represented as points, and relationships between the concepts are depicted as arrows (Axelrod, 1976, Hall, 1984). Statements such as 'the number of general practitioners has an impact on the amount of medicine prescribed by those general practitioners' and 'increasing the number of general practitioners may contribute to a reduction in costs', are represented in a cognitive graph as follows:

Figure 4.8: Causal relationships in a cognitive map

| | | |
|---|---|---|
| # of general practitioners | → | # medicine prescribed by general practitioners |
| # of general practitioners | → | costs of health care |

The cognitive mapping approach is used to transform the written statements of the pre- and posttest into a concept-relationship schematic by concentrating on the means-end and causal relationships included in the text. The advantage of this kind of reconstruction or reduction is that it is closely related to the way in which complex systems are represented in the model-building approach - both represent knowledge as variables (concepts) that are related to other variables by means of (causal) relationships. To assess the extent to which participants have changed their individual (mental) model or conceptualization of the problem, a 'model'-approach such as the cognitive mapping one, thus appears to be promising. Moreover, as stated by Vennix (1990, p. 116), the cognitive mapping approach is a powerful one, for not only does it allow for the inclusion of both means-end and causal relationships, it also is an excellent way to represent feedback processes, which are important in system dynamics thinking. Finally, it allows for the use of graph-analytical techniques to examine the properties of these cognitive maps.

Having described briefly what cognitive maps look like, account can be taken of the way in which the study's dependent variables have been operationalized on the basis of these cognitive maps. However, prior to that, a sample cognitive map[48] will be presented, on the basis of which the operationalization of each of the criteria will be illustrated.

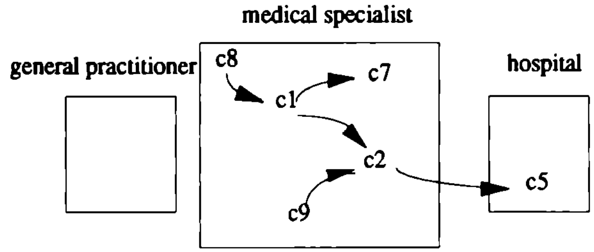
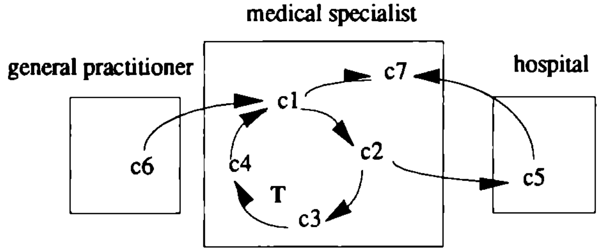
The cognitive maps depicted in figure 4.9, are supposed to be an answer to both the pre- and the posttest. It is an answer to question number two, concerning the subsystem of the medical specialist. To be able to interpret the cognitive maps, and to follow the procedures that will be described to measure the domain-specific and strategic knowledge components, additional information is added on the right-hand side of the two maps.

## Domain-specific knowledge

Regarding the domain-specific knowledge, an overview will be given of the way in which the subcriteria endogeneity, exogeneity, multi-disciplinarity, and knowledge of the external model incorporated in the conceptualizations, are operationalized in the present study on the basis of the pre-and posttest.

[48] The sample cognitive map is just a fictitious map, created by the author to illustrate the application of the operationalization outlined in the sections to come.

Figure 4.9: Sample cognitive map of both a pretest answer and a posttest answer to a question concerning the medical specialist subsystem

| | | domain-specific | strategic |
|---|---|---|---|
| pretest cognitive map |  | *endogenous*<br>concepts : c1, c2, c7, c8, c9<br>*exogenous*<br>concepts: c5<br>*monodisciplinary*<br>concepts: c7, c8, c9<br>*external model*<br>concepts c1,c2 | *feedbackloop: -*<br>*chaining* c8-c1-c7<br>c8-c1-c2-c5<br>c9-c2-c5<br>*connectivity*<br>relationships: 5<br>concepts: 6<br>*time-phase relationship: -* |
| posttest cognitive map |  | *endogenous*<br>concepts : c1, c2, c3, c4, c7<br>*exogenous*<br>concepts: c5,c6<br>*monodisciplinary*<br>concepts: c7<br>*external model*<br>concepts c1,c2,c4, c6 | *feedbackloop* c1-c2-c3-c4<br>*chaining*<br>(non-recursive)c6-c1-c2-c5-c7<br>c6-c1-c7<br>(recursive) c1-c2-c3-c4-c1<br>*connectivity*<br>relationships: 8<br>concepts: 7<br>*time-phase relationship*<br>c3-c4 |

1. Endogeneity and exogeneity

To measure the level of endogeneity and exogeneity included in a person's conceptualization (cognitive map), a distinction must be made between endogenous and exogenous concepts first. Endogenous concepts are concepts that, according to the researcher and on the basis of the model built by the participants, can be considered as belonging to the subsystem focused upon by the 'pre-and posttest' question. Concepts referring to the subsystem of the medical specialist, are regarded as endogenous, if the question concerns the medical specialist subsystem. Concepts pointing to one of the other two subsystems consequently are considered as exogenous concepts. With respect to the posttest map depicted in figure 4.9, it is clear that only the concepts 'C5' and 'C6' fall beyond the boundary of the medical specialist subsystem, and thus qualify as 'exogenous' concepts. Note that of these two exogenous concepts, concept 'C6' is also part of the external model.

Once the concepts of the cognitive maps have been classified as either endogenous or exogenous, the causal relationships between the concepts need to be classified as well. In order to qualify as endogenous, ***both*** concepts (cause ánd effect) of the relationship need to be of an endogenous nature. Exogenous relationships, by contrast, consist of cause-effect relationships that are made up of one or two exogenous concepts. If we apply this to the cognitive map of the pretest depicted in figure 4.9, the relationship between 'C1' and 'C2' for example, qualifies as endogenous, whereas the relationship between 'C2' and 'C5' will be considered as exogenous in nature. Note that to assess the endogeneity of the conceptualization, both the number of endogenous concepts and the number of endogenous relationships have been included to distinguish between maps with the same amount of concepts but different numbers of relationships among these concepts.

To illustrate the application of the procedures by means of which scores on the dimensions of endogeneity and exogeneity can be arrived at, look at the pretest scores on the endogeneity and exogeneity dimensions based on the *pretest* cognitive map depicted at the top of figure 4.9:

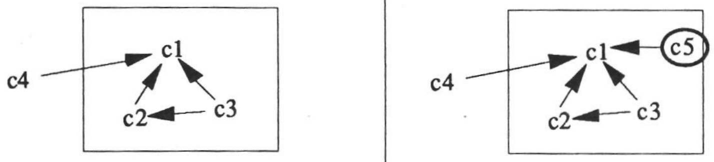
Figure 4.10: Endogeneity and exogeneity scores

| subcriterion | aspect | score |
|---|---|---|
| endogeneity | number of endogenous concepts | 5 |
| | number of endogenous relationships | 4 |
| exogeneity | number of exogenous concepts | 1 |
| | number of exogenous relationships | 1 |

The reason why it was decided to take the absolute number of concepts and relationships to assess the degree of endogeneity and exogeneity rather than the relative number of concepts and relationships (that is, divided by the total number of concepts and relationships), is because we are interested in real rather than relative improvement as far as the domain-specific content of the conceptualization is concerned. We are not primarily interested in, for example, a change from exogenous into endogenous concepts for this would be more a change in format than a change in content. To illustrate what is meant by an increase in the exogenous content of a conceptualization, look at the following two cognitive maps:

Figure 4.11: Exogeneity

*pretest* | *posttest*




The map on the left-hand side represents a possible answer to the pretest. It contains four concepts in total; three of them are endogenous and one of them is exogenous. Hence the relative exogeneity of the pretest answer is one-fourth (.25). The cognitive map on the right-hand side, depicting the same person's answer to the very same question, but this time to the posttest rather than the pretest, consists of five concepts rather than four, for 'C5' was added to the pretest answer. As a result of this increase in the number of concepts, a decrease in the relative exogeneity is brought about (.20), for the absolute number of exogenous variables has remained constant, that is, no exogenous concepts were added to the one already mentioned in the pretest. So if the relative exogeneity is employed to indicate how people do on the dimension of exogeneity, one does not measure the real knowledge participants have about exogenous factors, but only the amount of exogenous knowledge they have in relationship to their endogenous knowledge. In the realm of domain-specific knowledge, however, we want to know the amount of (endogenous or exogenous) variables (or relationships) one has knowledge about and consequently employs to account for the problem under discussion, rather than assessing how the proportions of endogeneity and exogeneity bear to one another.

2. Multi-disciplinarity

To measure the degree to which participants have included multi-disciplinarity in their conceptualization of the problem, again both concepts and relationships are being used. The procedure to determine the degree of multi-disciplinarity follows to a large extent the procedures carried out to assess the degree of endogeneity and exogeneity. One determines first as to whether a concept is multi-disciplinary in nature, and based on this, one assesses which of the individual relationships qualify as multi-disciplinary relationships. To classify as multi-disciplinary relationships, relationships must consist of at least *one* multi-disciplinary concept. Since it is extremely difficult for an outsider to determine as to whether a concept is multi-disciplinary or just mono-disciplinary in nature, it was decided to adopt a researcher-independent procedure to categorize the concepts into disciplinary and multi-disciplinary concepts. We simply decided to look at the way in which the concepts are used by both departments and determine on the basis of this whether a particular concept is department-specific or not.

According to this researcher-independent procedure, concepts qualify as (mono)disciplinary concepts if there is a *significant difference* in the degree to which they are used by the *two departments* at the *pre- and posttest*. Because there is a strict disciplinary distinction[49] between the two participating departments (one is a medical department, whereas the other is a economical-legal one), it is assumed that differences in departmental use of concepts can be used to assess differences in the disciplinary nature of the concepts. Once the concepts have been classified as either mono-disciplinary or

[49] The people who work for the medical department have taken a medical degree whereas the people from the other department have taken degrees in accounting, bookkeeping, and economics.

multi-disciplinary (those which do not qualify as mono-disciplinary concepts, automatically qualify as multi-disciplinary ones), an assessment can be made of the number of both kinds of concepts each of the participants uses to conceptualize the problem. So to arrive at such scores on the pretest, comparisons are being made between the two departments with respect to all the concepts used at the pretest. Based on this comparison, concepts are being classified. The very same procedure is repeated for the posttest, to arrive at a classification of concepts on the posttest as well.

Whether absolute or relative scores are to be used to assess the degree to which participants have increased the multi-disciplinary character of their conceptualization, the same line of argument as the one given in the discussion of the endogeneity and exogeneity criteria holds here: since we deal with domain-specific knowledge, our primary concern is the amount of multi-disciplinarity that is included in the cognitive map rather than the relationship between multi-disciplinary elements and the total size of the map.

Moreover, since it is expected that the number of concepts with a significant difference between departmental use is relatively low compared to the number of concepts where no significant difference is found, chances are that an increase in the number of concepts used in the posttest will bring about an increase in the multi-disciplinary nature of the representation - more concepts being equal to more multi-disciplinarity. This holds for both absolute and relative scores for it is the result of the way in which multi-disciplinarity is being defined (as the absence of significant differences in use). To illustrate the potential weakness of this definition of multi-disciplinarity, look at the following example of pre- and posttest scores with respect to the criterion of multi-disciplinarity:

Figure 4.12: Scores on multi-disciplinarity

| | pretest | posttest |
|---|---|---|
| number of concepts | 11 | 15 |
| number of multi-disciplinary concepts | 10 | 14 |
| number of mono-disciplinary concepts | 1 | 1 |

In this example, only one of the eleven concepts used at the pretest to describe the problem managed to classify as a mono-disciplinary concept, that is, a concept where significant differences in use between the two departments were found. As a result, the remaining ten concepts automatically qualified as multi-disciplinary ones, for the definition of multi-disciplinarity states that the absence of significant differences in use suffices to qualify as multi-disciplinary concepts. However, if this definition of multi-disciplinarity holds, increasing the size of the map (adding extra concepts) almost automatically results in an increase in multi-disciplinarity for it appears to be more likely to find and use concepts that are not specific for one department than to find or use concepts that are exclusively used by one of the two departments, as shown by the pretest scores. If we apply this definition of multi-disciplinarity to the present example, an increase in both absolute and relative multi-disciplinarity scores is found, despite of the fact that the number of concepts that qualified has not changed at all. To avoid such a potential contamination of increase in cognitive map and multi-disciplinarity (the risk of taking an increase in content for an increase of multi-disciplinarity), it was decided to use a somewhat more strict definition of multi-disciplinarity in which the number of concepts would not be incorporated. According to this strict definition, multi-disciplinarity can only be improved if one reduces the number of unique, department-dependent concepts to describe the problem at hand. Simply adding concepts which are not mono-disciplinary (and thus qualify as multi-disciplinary) is not enough. One has to make an effort to either change mono-disciplinary concepts for non-mono-disciplinary ones, or remove the mono-disciplinary ones from their conceptualization. As such, *an increase in multi-disciplinarity is defined in a 'negative' way, that is, as a decrease in mono-disciplinarity.*

As explained before, we do not want to use relative scores, for this can result in a situation in which the mono-disciplinary content of a cognitive map has not been changed, but its mono-disciplinarity score is changing due to a change in the total number of concepts used to describe the problem. However, it is felt that some kind of relativity needs to be taken into account for one would like to distinguish between mono-disciplinary concepts that are important in that they are used often, and concepts that only play a marginal role in the map, to arrive at a score of mono-disciplinarity that is doing justice to the person's conceptualization. Since mono-disciplinarity is considered as something that should be avoided (or at least be reduced), an assessment needs to be made of 'how bad the situation is', that is, how often these mono-disciplinary concepts are being used in the conceptualization; how important their role is in the explanation of the problem. Note that this line of reasoning does not apply to the criteria of endogeneity and exogeneity since there are no 'negative' elements in the cognitive map that should be removed as far as those two criteria are concerned. All elements (endogenous and exogenous) are considered to be valid and useful and by increasing the number of elements included in the conceptualization, an enrichment of one's conceptualization can be brought about. Hence it is suggested, in order to account for the relative importance of the concepts that have qualified as mono-disciplinary concepts, to include the number of relationships that are made up of mono-disciplinary concepts in the calculation of a multi-disciplinarity score. A higher number of mono-disciplinary relationships, indicating that the mono-disciplinary concepts are of greater importance, then would result in a higher score on the dimension of mono-disciplinarity and hence bring about a lower multi-disciplinarity score.

However, as shown in figure 4.12 the number of mono-disciplinary relationships has already been included in the operationalization of the multi-disciplinarity criterion. Rather than calculating a combined score on the level of the subcriterion 'multi-disciplinarity' directly, it was decided, for the time being, to distinguish between concepts and relationships, in line with the operationalization of the other criteria. The degree to which mono-disciplinary concepts are being used thus is taken into account in the measurement of multi-disciplinarity indirectly; it is accounted for by the separate aspect of mono-disciplinary relationships, measured by the number of mono-disciplinary relationships.

If we apply the above described procedure to determine the scores on multi-disciplinarity of the cognitive maps depicted in figure 4.9 (i.e. look at the absolute number of mono-disciplinary concepts and relationships as an indication of the multi-disciplinarity of one's cognitive map, a mono-disciplinary score of 3 regarding both concepts and relationships is arrived at at the pretest, while a score of 1 on the concepts and 2 on the relationships is found at the posttest. Since the posttest score is lower than the pretest score, this particular participant has indeed made an improvement on the dimension of multi-disciplinarity: a lower mono-disciplinary score tis regarded as an increase in multi-disciplinaryity.

3. External model

The last criterion to assess whether participants have changed their domain-specific knowledge, concerns the degree to which they have succeeded in incorporating the domain-specific knowledge contained in the external (conceptual and/or computer) model constructed and used during the sessions. To measure the degree to which the conceptualization of the problem is making use of the external model, a comparison is made between the concepts and relationships used by the individual participants and the concepts and relationships included in the external model. As such, the criterion resembles the Vennix (1990) epistemological criterion, in which the concepts and relationships of the cognitive maps of the participants are compared to the concepts and relationships of the external model as well. Note however, that due to the fact that the external model is not considered as an externally validated model (in contrast to the econometric model used by Vennix), qualifications such as validity or precision are not used in the present study - we refer to the degree of incorporation of the external model without any further qualification.

To qualify as a concept from the external model, the concept used by the participants has to be part of (or refer directly to) the external model. However, to qualify as a relationship from the external model, both the cause and effect concepts have to be included in the external model. Note that this is not to say that the relationships mentioned in the individual cognitive maps have in fact all been part of the external model, although almost all of them in fact were part of the external model. The reason why it was decided to consider relationships to be in agreement with the external model if both concepts (cause and effect) are elements of the external model, is to be able to deal with the fact that sometimes intermediate concepts are not included in a line of reasoning. To illustrate what is meant by this, consider the following causal chain included in the external model: 'A is bringing about a change in B, and B in turn, is having an affect on C'. As a result of this line of thought, a relationship exists between A and B, and B and C. If, for example, a participant decides not to include B in his or her causal explanation of a change in C and simply states that 'C is brought about by A', no correspondence would be found if the strict procedure was being applied according to which relationships only qualify if they literally exist in the external model. Our procedure, by contrast, acknowledges the fact that some kind of similarity exists between the two causal chains, despite of the fact that exact correspondence is lacking.

Regarding the issue of whether absolute or relative scores need to be used to assess changes in the knowledge people have (use) of the external model, the absolute number of concepts and relationships should be taken. This because the present study aims to assess how much knowledge the participants have about the external model rather than evaluating as to whether it may seem as if the use of elements of the external model is increased due to a decrease in the total number of concepts or vice versa. As such, the operationalization of this criterion does not differ from the previously discussed criteria.

To illustrate the procedure to arrive at a score on the criterion of external model, we return to the sample cognitive maps presented in figure 4.9. On the pretest, 2 concepts and 1 relationship managed to qualify as elements corresponding to the external model. On the posttest, however, 4 concepts and 3 relationships were found to qualify as corresponding to the external model. So with respect to this particular sample of cognitive maps, an increase in the number of elements 'borrowed' from the external model can be concluded to.

## Strategic knowledge

Having outlined how changes in the domain-specific content of the cognitive maps will be measured in the present study, account must be taken of the way in which the strategic knowledge component of the present study will be operationalized on the basis of the pre- and posttest. Hence, the present section will subsequently be concerned with the measurement of the criteria of feedbackloops, chaining, connectivity, and time-phase relationships.

1. <u>feedbackloops</u>

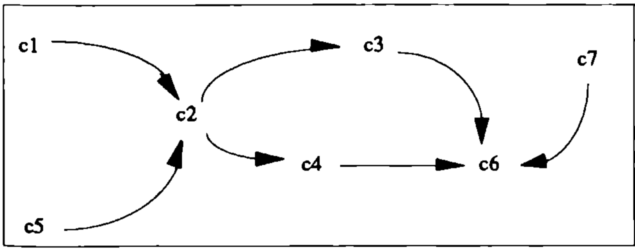
To measure the number of feedbackloops included in a conceptualization, all possible paths are examined (a path is a series of one or more consecutive causal relationships; a chain of causal relationships). Only if the very begin and end concept is identical, paths qualify as feedbackloops[50]. To measure the number of feedbackloops included in a cognitive map, one simply adds the number of paths that have qualified as feedbackloops. However, as indicated by Vennix (1990, p. 154), the number of concepts and the average length of the feedbackloops may affect the number of feedbackloops found using this particular method. This because the more concepts are included in a cognitive map, the more likely it is that paths exist that accidentally qualify as feedbackloops but have never been intended or explicitly stated as feedbackloops by the participants. The same holds, though in a somewhat different way, for the average length of the feedbackloop: the longer the average loop is, the more likely it is that an artificial increase in the number of feedbackloops is being brought about. Vennix consequently suggested to divide the number of feedbackloops found by the total number of concepts and the mean length of the feedbackloops. The present study by contrast, is somewhat hesitant to adopt this strategy for dealing with this problem of artificial or accidental qualification. First of all, it is felt that the Vennix solution, although technically sound, results in a difficult to interpret variable - the number of feedbackloops per concept per average length of feedbackloop. The second reason why it has been decided to postpone the decision as to whether the criterion of feedback has to be operationalized in a more complicated way than simply counting the number of feedbackloops, is because it remains to be seen as to whether indeed feedbackloops are constructed artificially by using the above described procedure. Hence it is decided to examine the relationship between the number of feedbackloops and the number of concepts and average length of feedbackloops first, before making a decision with respect to a possible adjustment of the current operationalization of the feedbackloop criterion.

Applying the provisional procedure to measure the criterion of feedbackloop outlined above, to the cognitive maps presented in figure 4.9, a score of 0 on the pretest and a score of 1 on the posttest is being arrived at. Consequently, it can be said that some improvements were made on the feedbackloop-dimension of the strategic knowledge represented in the person's cognitive map.

2. <u>chaining</u>

The second feature of system dynamics thinking that needs to be taken into account to evaluate the way in which the domain-specific knowledge is organized, is the so-called chaining. Recall that chaining was defined as the average length of the causal chains comprised in the cognitive maps. To measure the average length, both the length of the maximal paths and the length of the feedbackloops need to be calculated, for feedbackloops can be considered as a special kind of path. To arrive at the chaining score, the two subscores (length of maximal paths and length of feedbackloops) will have

[50] The starting point of a maximal path is the concept that itself is not being affected by any other concepts and thus only can serve as a cause concept. The ending point of a maximal path, is the concept which functions as an effect variable, without being a cause-variable itself. To illustrate, look at the following cognitive map in which five maximal paths can be discerned:




The following maximal paths can be discerned:
1) c1-c2-c3-c6
2) c1-c2-c4-c6
3) c5-c2-c3-c6
4) c5-c2-c4-c6
5) c7-c6

to be averaged. To illustrate this particular procedure, the chaining scores of the cognitive maps of figure 4.9 have been calculated. Regarding the pretest map, an average length of 2.33 was found, whereas a chaining score of 3.33 was arrived at with respect to the posttest map.

3. <u>connectivity</u>

The criterion of connectivity, defined as the degree to which the elements of the conceptualization are related to each other will be calculated by taking the ratio of relationships to concepts, that is, by dividing the total number of relationships by the total number of concepts included in a person's conceptualization. Note that this operationalization of connectivity coincides with the operationalization of the criterion of integration in the Vennix (1990) study.

Calculating the connectivity scores of the cognitive maps depicted in figure 4.9, a score of .83 on the pretest and a score of 1.14 on the posttest is being arrived at.

4. <u>time-phase relationships</u>

As explained before, the notions of amplification and delay have been combined to construct one criterion by means of which an assessment can be made of the degree to which participants are aware of the fact that the combination of feedback and delay can bring about unexpected dynamic effects to the behaviour of the system under consideration. To operationalize this construct, called time-phase relationship, explicit reference to time as an important factor in bringing about the dynamic behaviour of the system, is being looked for. So whenever the factor 'time' is included in the conceptualization of, say, the dynamic effects of a particular policy measure, a score on the time-phase relationship criterion will be awarded. The pretest cognitive map (figure 4.9). for example, does not have a score on the time-phase relationship criterion for no reference was being made to notion of time. In the posttest map however, time was being referred to once, as a result of which the posttest score on the time-phase relationship is equal to 1.

The operationalization of the dependent variables and their application to the sample cognitive maps (pre- and posttest) depicted in figure 4.9, can be summarized as follows:

Figure 4.13: Outcomes corresponding to the maps depicted in figure 4.9

| *Individual and inter-individual change* | | | score | |
|---|---|---|---|---|
| main criterion | subcriterion | aspect | pretest | posttest |
| domain-specific (II and IV) | endogenous | concepts | 5 | 5 |
| | | relationships | 4 | 5 |
| | exogenous | concepts | 1 | 2 |
| | | relationships | 1 | 3 |
| | multi-disciplinary | concepts | 3 | 1 |
| | | relationships | 3 | 2 |
| | external model | concepts | 2 | 4 |
| | | relationships | 1 | 3 |
| strategic (I and III) | structure | feedbackloop | 0 | 1 |
| | | chaining | 3.33 | 4.33 |
| | | connectivity | .83 | 1.14 |
| | amplification | time-phase relationship | 0 | 1 |
| | delay | | | |

## 4.5.2 OPERATIONALIZATION OF THE POTENTIALLY SPECIFYING VARIABLES

The next step in the operationalization of the present study's theoretical model concerns the potentially specifying variables. As depicted in figure 3.11, background characteristics, evaluation of the process, and time-investment are considered to be important in that they may specify the effects of the independent variable on the dependent one, that is, specify the impact participative policy modelling is having on the conceptualization of the problem. To start with, an overview will be given of the operationalization of the background variables included in the present study. This will be followed by the operationalization of the process variables that represent how the participants feel about the program in terms of components and aspects. Finally, a description is given of the operationalization of the time-involvement variable.

### *Background characteristics*

The background variables included in the present study are age, educational background, area of expertise (specialism), number of years with the firm, policy experience, department, gender, and organization. They have been measured as follows (with the exception of the variable 'organization'):

Figure 4.14: Operationalization of the backgroundvariables

| variable | categories |
|---|---|
| age (interval ) | (20-29) (30-39) (40-49) (50-59) (60-older) |
| education (ordinal) | (secondary education) (intermediate vocational education) (higher vocational education) (university) |
| area of expertise (nominal) (specialism) | (general practitioner) (medical specialist) (hospital) (no particular area) (general practitioner + medical specialist) (medical specialist + hospital) (general practitioner + hospital) |
| years with the firm (interval) | (0-1) (2-3) (4-5) (6-10) (more than 10) |
| policy experience (interval) | (> 80 % of time) (60-79 % of time) (40-59 % of time) (20-39 % of time) (< 20 % of time) |
| department (nominal) | (financial-legal) (medical) (other) |
| gender (nominal) | (male) (female) |
| organization (nominal) | (organization 1) (organization 2) |

Regarding the variable *age*, since it was known beforehand that none of the participants would be younger than 20 years of age, only categories representing 20 years of age and higher were included in the questionnaire.

*Education* can be considered as an ordinal variable, for it represents a rising number of years of formal education taken by the participants. However, because the participants were also asked to state the exact kind of education they had been taking, a nominal aspect of the variable of education can be discerned as well.

*Area of expertise* is based on the subsystems that can be distinguished within the Dutch Health Care system. Recall that a general practitioner subsystem, a medical specialist subsystem, and a hospital subsystem were discerned. The answering categories thus are a combination of these subsystems, whereby a combination of all three is represented by the fourth box, saying 'no area of expertise', indicating that the person is knowledgeable in all three fields rather than not being knowledgeable at all.

*Policy experience* is measured by asking the participants how much of their working time is spent on policy making issues in the organization.

Generally speaking, the present study's participants come from two *departments*: a economical-legal department called 'Overeenkomst en Tarieven' and a medical department called 'Medische Dienst'. The reason why, however, it was decided to add a third category ('other'), was because it was known that the head of the two departments, the Chief Executive Officer (CEO), would be participating as well.

The background variable '*organizational background*' referring to the pre-merger organization the participants worked at prior to the merger process, was not included in the questionnaire because the information could easily be provided for by our liaison person.

### *Evaluation of the process*

Having outlined the way in which the background variables have been measured in the present study, account must be taken of the operationalization of the second category of potentially specifying variables, the variables that are concerned with the evaluation of the process, that is the participative policy modelling sessions.

To start the description of the operationalization of the process variables, an overview will be given of the way in which participants feel about the three major components of the participative policy modelling method: preparatory texts (so-called workbooks), small-group activities, and plenary presentation and discussion. Following the operationalization of the major components, a presentation will be given of the way in which aspects such as usefulness and easiness have been operationalized in the questionnaire. Finally, two miscellaneous items, which do not fit into any particular aspect or component, will be discussed briefly.

### *Major components*

#### 1. preparatory text (workbook)

Participants were asked to read three workbooks to prepare for the first three sessions. Each of the workbooks was handed out about two weeks before the session itself was taking place, to make sure that the participants had sufficient time to read the workbooks and complete the exercises that were included in it. They were given only one workbook at a time to prevent them from reading ahead and finish all three workbooks before any of the sessions had taken place.

It was estimated that to read the first workbook, providing a general introduction to the model-building process, and outlining the preliminary health care model, approximately three hours would be required. The second and third workbooks, regarding 'external influences' and 'policy measures', by contrast, were expected to take one hour at a maximum each to read. Hence, it was expected that, on the average, about five hours of homework were required to successfully take part in the participative policy modelling sessions. As to whether participants indeed did take that much time to prepare, will be shown in Chapter Six, concerning the results of the present study.

To measure the degree to which the participants felt positively about the preparatory text, the following nine questions were included in the questionnaire (the numbers in front of the questions refer to the numbers they have in the questionnaire, as depicted in Appendices 1 to 4):

Figure 4.15: Operationalization of the workbook

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. Reading the workbook was | easy | 5 | 4 | 3 | 2 | 1 | difficult |
| 2. The workbook was | interesting | 5 | 4 | 3 | 2 | 1 | uninteresting |
| 3. From reading the workbook I have | learned a lot | 5 | 4 | 3 | 2 | 1 | learned nothing at all |
| 4. Reading the workbook was | useful | 5 | 4 | 3 | 2 | 1 | useless |
| 5. I consider the workbook to be | too long | 1 | 3 | 4 | 5 | 2 | too short |
| 6. To prepare for today's session the workbook was | good | 5 | 4 | 3 | 2 | 1 | bad |
| 9. The exercises in the workbook were | clear | 5 | 4 | 3 | 2 | 1 | not clear |
| 10. The exercises in the workbook were | easy | 5 | 4 | 3 | 2 | 1 | difficult |
| 11. The exercises in the workbook were | useful | 5 | 4 | 3 | 2 | 1 | useless |

Note that all nine question are of an ordinal level. Except for question number five, the pattern of points attributed to each of the answering categories goes from 5 to 1, that is, 5 points are given to the most positive answer (extreme left-hand side), whereas 1 point is given for the most negative one (extreme right-hand side). Recall that scores lower than 3 indicate that the person did not appreciate the aspect of the workbook very much whereas a score higher than 3 shows that the person is feeling positively about the very same aspect.

With respect to question number 5 ('*I consider the workbook to be too long versus too short*'), it is clear that neither of the two extremes can be used to express a positive view on the workbook. Hence it was decided to recode the answer to arrive at answering categories ranking from 1 to 5, whereby 1 point is given to the least positive qualification of the workbook, and 5 point are given to the most positive one.

Since it was felt that for a preparatory text, being too long is far more negative than being too short, the box on the extreme left was given one point. However, being too short can be regarded as a disqualification as well, for the workbook should have been longer then. Hence, the extreme right box was ranked second, that is, was awarded two points. The box right of the middle is considered as the most positive answering category for it is felt that people who say that the workbook has been somewhat short (rather than saying too short), have enjoyed reading the workbook and would not have minded spending some more time reading. As a result, five points are given to this box. With respect to the remaining two boxes (the one left of the middle and the middle one itself), the middle one is considered as being more positive than the one left from the middle indicating that the workbook had been somewhat long. As a consequence, four points are attributed to the box in the middle, whereas three points are given to the one left from the middle.

This concludes the description of the operationalization of the evaluation of the preparatory text. Next, the small-group activities will be focused upon.

## 2. Small-group activities

One of the other major components of the participative policy modelling program consists of activities that are carried out in small groups. During these small-group sessions, participants are provided with the opportunity to interact intensively with both other participants and the computer model under construction. To assess how the participants felt about this particular component of the program, the following six questions were included in the questionnaire:

Figure 4.16: Operationalization of the small-group activities

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 12. The assignments and discussions of the small-group session were | easy | 5 | 4 | 3 | 2 | 1 | difficult |
| 13. The assignments and discussions of the small-group session were | interesting | 5 | 4 | 3 | 2 | 1 | uninteresting |
| 14. From the assignments and discussions from the small-group session I have | learned a lot | 5 | 4 | 3 | 2 | 1 | learned nothing at all |
| 15. The assignments and discussions of the small-group session were | useful | 5 | 4 | 3 | 2 | 1 | useless |
| 16. The amount of time spent on the assignments and discussions of the small-group session was | too long | 1 | 3 | 4 | 5 | 2 | too short |
| 17. The assignments and discussions of the small-group session have increased my knowledge of how my colleagues think about the health care system | agree | 5 | 4 | 3 | 2 | 1 | disagree |

Question number sixteen has been recoded for the same reason and in the same way as question number five. Both concern the amount of time that is spent on a particular component of the participative policy modelling method.

## 3. Plenary presentation and discussion

Following the small-group activities, plenary presentations and discussions were held to present and discuss the outcomes with the rest of the participants. Each of the subgroups (most of time three per session) was given the opportunity to present and explain the outcomes of these activities to the other participants. Based on this presentation, a discussion was started in an attempt to arrive at a more shared perspective on the issues at hand.

To evaluate how the participants felt about this particular component of the program, they were asked to answer the following four questions:

| Figure 4.17: Operationalization of the plenary session | | | | | | | |
|---|---|---|---|---|---|---|---|
| 18. The plenary presentation and discussion was | interesting | 5 | 4 | 3 | 2 | 1 | uninteresting |
| 19. From the plenary presentation and discussion, I have | learned a lot | 5 | 4 | 3 | 2 | 1 | learned nothing at all |
| 20. The plenary presentation and discussion was | useful | 5 | 4 | 3 | 2 | 1 | useless |
| 21. The amount of time spent on the plenary presentation and discussion was | too long | 1 | 3 | 4 | 5 | 2 | too short |

It is shown in figure 4.17 that the more one has learned from the component, the more interestingly and useful one has found the plenary component, and the more appropriate one feels the time spent on it has been, the more positive the evaluation of this component of the program will be.

Note that question number 21, concerning the amount of time spent on the component, has been recoded, in line with the way in which the questions number sixteen and five were recoded.

## *Major aspects*

The second way to evaluate the program is to focus on aspects or qualifiers rather than on components that make up the participative policy modelling method. By concentrating on these qualifiers, no distinction is being made between any of the components that make up the program, as a result of which the scores that are being arrived at, concerning a particular aspect or qualifier, evaluate the program in its totality. As stated previously, the aspects that will be used for this concern the question of whether the program was easy, interesting and useful, whether one has learned something from it, and finally, whether the right amount of time was spent on it. Since the very same items have already been described in the operationalization of the three major components, it suffices to simply present the numbers of the individual items by means of which a score on each of the five aspects can be arrived at (for an overview of the questionnaire, the reader is referred to Appendices 1 to 4):

| Figure 4.18: Aspects/qualifiers | |
|---|---|
| easiness | # 1: reading the workbook was easy vs difficult<br># 10: the exercises in the workbook were easy vs difficult<br># 12: the assignments and discussions of the small-group session were easy vs difficult<br># 36: the fourth session was easy vs difficult |
| usefulness | # 4: reading the workbook was useful vs useless<br># 11: the exercises in the workbook were useful vs useless<br># 15: the assignments and discussions of the small-group session were useful vs useless<br># 20: the plenary presentation and discussion was useful vs useless<br># 39: the fourth session was useful vs useless |

| Figure 4.18: Aspects/qualifiers (cont.) | |
|---|---|
| interesting | # 2: the workbook was interesting vs uninteresting<br># 13: the assignments and discussions of the small-group session were interesting vs uninteresting<br># 18: the plenary presentation and discussion was interesting vs uninteresting<br># 37: the fourth session was interesting vs uninteresting |
| learned from it | # 3: from reading the workbook I have learned a lot vs learned nothing at all<br># 14: from the assignments and discussions of the small-group session I have learned a lot vs nothing at all<br># 19: from the plenary presentation and discussion I have learned a lot vs nothing at all<br># 23: altogether, the session has resulted in an increase in my knowledge about the health care system agree vs disagree<br># 39: from the fourth session I have learned a lot vs nothing at all<br># 41: the fourth session has increased my knowledge agree vs disagree |
| time spent | # 5: I consider the workbook to be too long vs too short<br># 16: the amount of time spent on the assignments and discussions of the small-group session was too long vs too short<br># 21: the amount of time spent on the plenary presentation and discussion was too long vs too short |

Having outlined the way in which the evaluation of the program has been operationalized in the present study, two more or less miscellaneous items need to be dealt with. The first items concerns the degree to which people feel that the workbook they have been reading is linked up with the way in which they deal with the health care system in their daily work (question number 6). Since one can argue both that being linked up with and not being linked up with their daily work can be considered as being positive, for one either has the opportunity to relate the information to the person's world of work, or is being provided with a refreshing new point of view, it was decided not to include this item in the list of items measuring the 'quality' of the workbook. The other item that has not been mentioned so far, concerns the relationship between the various components (question number 22). It is only when a score for the program in its totality rather than its components is being calculated, that this item will have to be taken into account.

### *Tme-involvement*

Finally, a description needs to be given of the way in which time-involvement has been operationalized. To determine the participants' time involvement, both attendance of the sessions, and preparation are to be taken into account. Because it is known who attended which session, and we also know how long each of the sessions took, a score on attendance can easily be arrived at in terms of hours of participation. To measure the amount of time spent on preparation, participants were asked to answer question number eight of the questionnaire. Notice that in contrast to Vennix (1990), it was decided not to ask the participants how thoroughly they had been reading the preparatory text, for it was felt that this would be an inappropriate thing to do.

Now that a description has been given of the way in which the dependent variables and potentially specifying variables have been measured in the present study, we will once more focus on the stimulus or treatment (i.e. the participative policy modelling program) that is playing such a central role in the present study. The description of the program will be supplementary because the program has, to some extent, already been described in Chapter Two. Since the activities carried out in the participative policy modelling sessions have already been described in Chapter Two (for a summary, the reader is referred to figure 2.12), the following description of the program will concentrate on the program's time-table, to provide additional information on who is doing what and when.

## 4.6 PARTICIPATIVE POLICY MODELLING: TIME-TABLE

Having described the present study's design, its research methods, and the way in which its dependent and potentially specifying variables have been operationalized, account needs to be taken of the actual time-schedule that was used to supplement the description of the study's stimulus or treatment (i.e. the participative policy modelling method) of Chapter Two. It is important to focus on the time-table to illustrate that in spite of the fact that three of the four sessions were carried out in two separate groups rather than in one, the two subgroups can be considered as one and the same experimental group. As already explained, the program offered to the two groups was exactly identical, and the outcomes of the group that was first to participate in a particular session (note that the order was alternated wherever possible), served as additional input for the other group. Moreover, the computer model that was eventually arrived at, was the same for both groups. The description of the time-schedule, thus will be used to provide additional support for the assumption that the two groups have in fact been participating to the same program both in terms of content and time at which the activities were carried out. Besides, it will show how the activities of the present study's research and the participative policy modelling are related to each other with respect to the dimension of time.

Figure 4.19: Time-table

| activity | group I | group II | time |
|---|---|---|---|
| introduction | sept. 20th 1990 | sept. 20th 1990 | .5 hours |
| handing out workbook #1 | sept. 24th | sept. 24th | |
| reading workbook #1 | sept. 24th-oct. 2nd | sept. 24th-oct. 2nd | 3 hours |
| session # 1 | oct. 9th | oct. 10th | 3 hours |
| handing out workbook #2 | oct. 9th | oct. 10th | |
| reading workbook #2 | oct. 9th-nov. 2nd | oct. 10th-nov. 2nd | 1 hour |
| session # 2 | nov 20th | nov 22th | 3 hours |
| handing out workbook #3 | nov. 20th | nov. 22th | |
| reading workbook #3 | nov. 20th-nov. 29th | nov. 22th-nov. 29th | 1 hour |
| session # 3 | dec. 18th | dec. 20th | 3 hours |
| session #4 | jan. 22nd 1991 | jan. 22nd 1991 | 2 hours |
| evaluation | jan. 22nd-jan 29th | jan. 22nd-jan 29th | .5 hours |

As depicted in figure 4.19, the lead time of the study almost equalled six months. The first activity (the pretest) was carried out in September 1990, while the closing posttest was completed just before the end of January of the following year.

Regarding the amount of time reserved for the pre- and posttest, we deliberately put 30 minutes only on the time-schedule handed out to the participants, to ensure that the participants would not be put off reading that they had to do something prior and posterior to the actual participative policy modelling sessions. This is also the reason why the pretest was called 'introduction' rather than pretest. The second reason why participants were told that the pre- and posttest would take about 30 minutes each, was to avoid them from putting in too much time in trying to come up with a 'too good' answer to the questions. We did not want them to, say, go to the library and read some articles on how to reduce the costs of health care. By contrast, since what we were after was the amount of knowledge they have at hand, ready to use, it was felt that not too much time should be taken to gather information before stating an answer to each of the three questions of the pre- and posttest.

Wherever possible (it had to fit in the participants' agendas), the order at which the two groups participated in the sessions alternated. However, each of the pairs of sessions (that is, the very same session for group 1 and group 2), was conducted in the same week, to avoid differences (e.g. differing external influences) between the two groups.

The participants were asked to read the workbook a relatively short time before the actual session was carried out to ensure that the relationship between preparation (workbook) and actual session could not be ignored. Moreover, having the participants complete their workbooks just prior to the sessions, increases the likelihood of having them remember the content of the workbooks and the answers they gave to exercises included in the workbooks, so that they will be better prepared. We obviously had to allow for at least one week in between the completion of the workbook and the session, to be in the position to read the workbooks and use them to prepare for the next session.

Finally, by summing the time estimated to complete the various activities of the participative policy modelling program, an estimate of the time required for individual participants to participate in the program, as well as the total time to be invested by the client organization, can be acquired. For individual participants, full participation was estimated to take about 17 hours, including the one hour that was scheduled for completing the pre- and posttest. Since 31 people were taking part in the participative policy modelling program, an investment of about 530 hours in total had to be made by the client organization.

## 4.7 SUMMARY

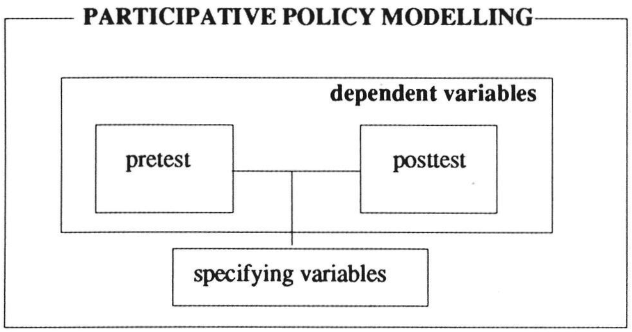
The current chapter has been concerned with some important methodological issues regarding the *empirical* evaluation of the effects of participative policy modelling on the way in which those who take part in it look upon a particular policy problem. To start with, the implications of the notions of 'historicism' and 'situationalism' were discussed, for it was felt that the status of the knowledge being arrived at during the policy making sessions, would apply to the outcomes of the present study as well.

Following the discussion of these implications, the distinction between intended and actual design was introduced. This to illustrate the kind of adjustments we had to make to the design we originally had in mind, in order to live up to the participating organization's requirements. It was argued that, although the actual design can be regarded as a somewhat weaker design that the intended design (the loss of the use of a control group can be considered as a serious blow to the present study's external validity), most of the potential threats to the study's internal validity could be made implausible.

Next, the methods employed in the present study were focused upon. An overview was given of the questionnaire, followed by a description of the pre- and posttest used in the present study to assess the effects of participative policy modelling.

In the sections that followed, an overview was given of the way in which the present study's variables are operationalized. To start with, the dependent variables were focused upon. Following the description of the way in which these variables are operationalized on the basis of both the questionnaire and the pre- and posttest, the operationalization of the present study's potentially specifying variables was concentrated upon. The present study's theoretical model can be depicted as follows:

Figure 4.20: Theoretical model




Finally, a more detailed description of the present study's independent variable (i.e. the treatment, the participative policy modelling program) was given in terms of who was doing what and when.

In the next chapter, the preliminary model of the Dutch Health Care System will be examined in detail. This to illustrate both the complexity of the system our policy makers were taking part in, and to provide some understanding of the activities that are taking place in participative policy modelling sessions.

# CHAPTER 5: THE DUTCH HEALTH CARE SYSTEM

## 5.1 INTRODUCTION

Having described the research design, the stimulus of which the effect will be evaluated, and the way in which these effects will be measured, account must be given of the organization participating in the present study. Not only do we have to describe the organization itself and the people who participated in the sessions, it is important to outline the reasons why the organization was willing to participate as well. Bear in mind that not too many organizations have been willing to take part in a participative policy modelling program and spend extra time on research-based activities such as questionnaires or pre- and posttests. The limited number of reports on successful real-life evaluations serve as an indication of the difficulties researchers have faced in the attempt to evaluate the effects of participative policy modelling methods in a real-life situation rather than by means of laboratory experiments.[51] Prior to describing the reason(s) why the organization was willing to take part in the program, a description of the changes that were expected to take place in the Dutch Health Care System in general and the health care market in particular at the time of our study, will be given. This because the reason why the organization was willing to take part in the program can be regarded as an attempt to prepare for some of these changes.

Following the description of the developments that have taken place in the Dutch Health Care System, a presentation will be given of what regional health care insurance organizations in the Netherlands are in general (e.g. what their objectives are, how they are financed), to arrive at an understanding of the impact the expected changes in the health care system may have on our client, a regional health care insurance organization.

Once the major characteristics of regional health care insurance companies have been described, the organization that was willing to participate will be focused upon. In the description of the organization some of the recent developments within the organization (the organizations had just completed a merger process) and the departments that participated in the program will be examined.

Finally, an overview will be given of the preliminary model of the Dutch Health Care System presented to the participants in the first workbook to provide an idea of the structure and dynamics of the health care system in which regional health care insurance organizations operate, and to illustrate what the preliminary model discussed by the participants looked like.

## 5.2 BRIEF DESCRIPTION OF THE DUTCH HEALTH CARE SYSTEM

To understand the role regional health care insurance organizations play in the Dutch Health Care System, and to arrive at an understanding of why our health care insurance organization was willing to participate in the present study, a brief description of the Dutch Health Care System will be given to start with. In this description, the dualistic nature of the system will be discussed first under the heading of 'halfway competitive markets'. Next, the issue of rising costs of health care will be touched upon, leading to a discussion of the changes suggested by the Dekker-committee in the second half of the 80s, to deal with this rise in costs. To examine the effects that measures proposed by the Dekker-committee[52] may have on regional health care insurance organizations, a

---

[51] Difficulties have for instance been reported by Vennix (1990) and Gould (1989a).

[52] Since a new government has come to power (1990), the plans concerning the Dutch Health Care System have been given a different name. They are now called after the present-day's Parliamentary Under-Secretary

presentation will be given of what the responsibilities and activities of regional health insurance organizations in the Netherlands are, and why and when these regional health care insurance organizations were established. Moreover, the presentation of the regional health care insurance companies will include an overview of some of the current developments that have taken place (or are about to take place) 'in the land of regional health insurance companies'. Note that these developments refer to changes that have taken place till the year 1990.

### 5.2.1 HALFWAY COMPETITIVE MARKETS

The Dutch health care system can roughly be characterized as a system that is located somewhere between the fully nationalized British National Health Service (NHS) and the market-based U.S. health care system. This means that besides partial competition through price mechanisms and the size of a policy, and partly regulated competition between non-profit organizations, planning institutions consisting of (among others) representatives of employers, employees, health care insurance organizations, and federal government play an important role in the health care system.

Within the context of this decision-making structure, two important decisions were made in the past, that have affected the character of the Dutch health care system deeply.

The first decision concerns the implementation of a mandatory insurance, according to which individuals, with an annual income below 50.000 Dutch guilders are automatically and obligatory insured against the costs of health care[53]. About sixty per cent of the Dutch population falls within this category and is consequently automatically insured against health care costs. Their health care consumption is taken care of (both financially and administratively) by non-profit regional health care insurance companies that have a social character and are based on governmental regulations. These regional health care insurance companies are called 'ziekenfondsen'[54]. How this system of mandatory insurance is financed, i.e. where the 'ziekenfondsen' get their money from, will be explained in more detail in the next section concerning the Dutch 'ziekenfondsen' in general. Individuals who have an annual income higher than 50.000 guilders, can (they certainly do not have to) insure themselves against the costs of medical health care at one of the many national (rather than regional) private health care insurance organizations that work on a commercial basis (these commercial health care insurance organizations can be profit or non-profit in nature).

The second important decision dates back to 1974 (Structuurnota, 1974). It concerns the division of the health care system into echelons. In order to be able to look at the health care system more systematically, and be more successful in planning the health care system in an attempt to contain its costs, services that could be regarded as functionally similar (from a patient's point of view), were put together into one and the same echelon. As a consequence, a distinction was being made between general health care services (e.g. practice- and community nurses, health care workers, and family

---

of Health Care, Mr. Simons, who's from the Labour Party. One of the most important differences between the plan proposed by the Dekker-Committee in 1987 and the plan developed by Mr. Simons, concerns the size of the mandatory package for which every Dutchmen is automatically insured. In the Dekker proposal, about 85 per cent of all services were part of this National Health Insurance package. The remaining 15 per cent had to be insured on a voluntary basis. In the plans developed by Simons however, about 3 percent of all services would be excluded from the basic package as a result of which many have felt that the idea of market-driven costs-containment being a central element of the Dekker-proposal had been given up. Partially due to this change in plans, some commercial insurance companies have gradually withdrawn themselves from the health care insurance market. Those commercial health care insurance organizations that have remained on the health insurance market, have almost all formed strategic alliances with one or more non-commercial health insurance companies (the regional health care insurance organization which no longer work purely on a regional basis).

53 If married, dependents are insured as well.

54 The regional health care insurance organization participating in the project is such a 'ziekenfonds'.

physicians/general practitioners) and more specialized health care, provided by medical specialists[55]. Note that in the Dutch health care system patients have no direct access to the second echelon for consultation and treatment except for emergencies (e.g. car accidents, heart-attacks). Patients have to see their general practitioner first in order to be referred to a medical specialist. As will be explained in more detail later on in the description of the model of the Dutch health care system, specialists can refer the patient back to his or her general practitioner for continuation of the medical treatment if necessary, or release the patient from further medical treatment.

Now that it has been explained why the Dutch health care system can be considered as a halfway commercial market system, and an overview has been given of the basic structure of the system in terms of echelons[56], account needs to be taken of the dynamics of the system both in terms of costs and policy measures recently taken (in 1987 and 1988 the plans for these changes were made, as will be explained in more detail in the next section) by the Dutch federal government to contain the costs of health care. The description of the development of costs and the changing role of government will make it clear why the regional health care insurance organization taking part in the present study had felt the need to participate in a model-building process focusing on the costs of health care.

### 5.2.2 THE COSTS OF HEALTH CARE

The costs of health care have risen dramatically over the last few decades. The total health care costs increased from 3.9 per cent of the gross National Product in 1968 to about 10 per cent in 1985 (Bex, 1989b). The Dutch Health Care System is not unique in this respect. Comparisons to other civilized countries show that almost everywhere the same increase in costs of health care has taken place (Schieber & Poullier, 1988). Several causes for this gradual but persistent increase have been identified (Grünwald, 1987; Honigh, 1983; Van der Maessen, 1987; Montfort & Vandermeulen, 1988). However, most of them are exogenous in nature; they affect the health care system but are not affected by the system themselves. Examples of such exogenous factors are increasing wage rates and increasing energy prices. However, as Vennix, Gubbels, Post, and Poppen (1988) put it:

> "Up to now, very little attention has been devoted to the internal dynamics of the health care system, which could also be held responsible for cost increases. In addition, most policy options aimed at cost reduction do not take these internal dynamics into consideration, though they could possibly neutralize the expected effects of a policy option."

Examples of policy measures that have not been too successful, possibly because of the 'unpredictable' dynamic behaviour of the Dutch Health Care System, are 'the introduction of a copayment system for treatment by medical specialists and use of medicine' to make the consumers more cost-sensitive, so that the utilization of these services will be reduced' (Schrijver, 1987), and 'the decrease of the utilization of the more expensive second and third echelons by increasing the first echelon, that is, the number of family physicians'[57] (Honigh, 1983; Kroonen & Bex, 1989).

---

55 Medical specialists can be found in both hospital settings and in facilities for long-term patient care.

56 For a more detailed description of the history and structure of the Dutch Health Care System, the reader is referred to Boot and Knapen, 1984

57 In Dutch, these two measures are called 'specialisten-geeltje' and 'medicijnenknaak' Post (1984a) for instance, describes how the introduction of a (fixed-guilder) copayment system for medicine (the so-called 'medicijnenknaak') resulted in a change in prescription from the less expensive to the more expensive medicines, and an increase in the amount of medicine prescribed per recipe As a result of this the costs of medicine showed an increase rather than a decrease.

Because the Dutch government had become increasingly concerned about the continuous rise in costs of health care and the lack of success policy measures have had so far, a special committee[58] was formed to think about changes that could be made to the health care system to control the rate of growth of health care spending. The committee was asked to develop plans to control the costs of health care by increasing the competitiveness of the system. In other words, the committee was asked to look for ways to control the costs within the context of withdrawing federal government and an increase in negotiations between the providers and financiers of health care (Schrijver, 1987). In 1987, the committee's report 'Bereidheid tot verandering [Willingness to Change]' was sent to parliament for discussion. Based on this, the Lubbers II-cabinet (1988) announced a series of structural changes to the health care system in a report called 'Nota verandering verzekerd' [Change insured].

The changes proposed in this report that seem most important to understand the changing role of regional health care insurance organizations in general (and our client organization in particular), can be summarized as follows:

* Rather than a government-led health care system, the system will be changed into a market-based one. It is expected that this, due to an increase in the negotiations between providing and financing agencies, will result in an increase in the system's efficiency (Roscam-Abbing, 1989; Schrijver, 1987; Ven, van de, 1989; Wijnberg, 1989)

* To change the Dutch Health Care System into a market-based system (but still based on a social law), the existing health care *insurance* system will have to be changed. The existing system, in which a distinction is being made between mandatory coverage (premium is based on the individual's income and is paid by both employee and employer) and private health care insurance coverage (premium is independent of the individual's income; a so-called nominal sum), will have to be replaced by a system where one and the same basic benefit package, covering 85 per cent of all services, is offered to everybody as if it were a true national health care system. On top of this basic package, additional coverage (covering the remaining 15 per cent) can be bought on a voluntary basis. The premium for the basic package is paid for by both employer and employee (as part of the tax premium), and depends on the individual's income. Health insurance companies cannot vary the premiums people have to pay for the basic package. Hence competition on the basis of this basic package can be ruled out. The premium paid for additional coverage however, is not related to the individual's income, and insurance companies are allowed to set their own premiums. Note that the existing difference between mandatory and private health care insurance companies will disappear once the Dekker-proposals have been introduced.

  Much is expected of the additional package in the endeavours to reduce the costs of health care, for it is believed that the additional package can serve as an incentive to work in an efficient, cost-effective manner. This because the more efficient one is (that is, the cheaper the services are that are bought from the providers of care), the more competitive one can be for it enables insurance companies to lower the premium they ask for their additional package. It is expected that negotiations between providers and financiers (insurance companies) about the price and quality of the services provided thus will become increasingly important in the Dutch Health Care System (Schellekens, 1989; Zitman, 1989).

* In order to be able to negotiate about the price and quality of care, health care insurance organizations no longer are required to sign contracts with all providers of

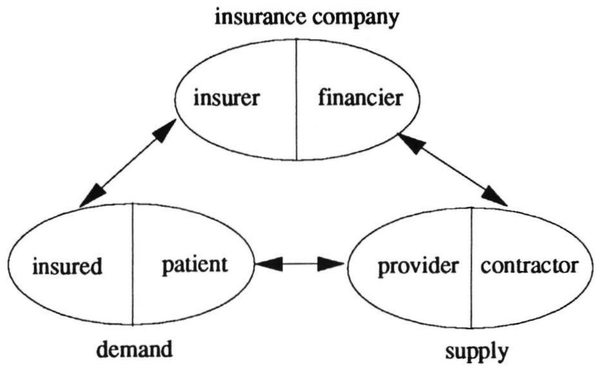
---

[58] The committee was called 'de commissie structuur en financiering van de gezondheidszorg' ['the structure and finance of the health care system' committee], also called after its chair Dr. Dekker, the 'Dekker-committee'. The committee received its assignment in March 1987.

health care. They are given the liberty to exclude providers (e.g. family doctors, hospitals, or physicians/medical specialists) who are, say, too expensive, from their insurance program. As a consequence, the services offered by those providers who are excluded from the program are not covered by the insurance company (Roscam-Abbing, 1989).

* Moreover, it is suggested that three major parties (stakeholders) will remain and that consequently three major markets can be distinguished in the Dutch Health Care System to be (Roscam-Abbing, 1989). The relationships between these three parties can be represented as follows (Post, 1989):

Figure 5.1: Dutch Health Care System: major parties




Note that each of the three parties is playing two roles at the same time, thereby contributing to the complex, difficult to predict character of the health care system.

The demand-side of the health care system consists of people who can be considered as both patients and insured individuals. As patients, they do not care about the price of the services and are only concerned with getting the appropriate services. As policy-holders however, that is, as insured individuals, they are very much price-oriented for they want to limit the premium they have to pay. Hence, a potential conflict of interest within this group of stakeholders can be found.

The supply-side of the health care system has a conflict of interest as well. On the one hand, as providers of health care, suppliers want to serve their patients' needs as much as possible, that is, they have the interest of their patients at heart. As contractors, by contrast, they are concerned about the costs that are involved because they do not want to jeopardize their relationship with their financiers, that is, the insurance companies that worry about profitability. In addition to this, suppliers of health care earn a living by 'selling' services. This obviously can be an incentive to 'sell' more and increase the price of their products/services.

As far as the insurance companies are concerned, duality is brought about by the roles of insurer and financier of health care that insurance companies play. As insurers, they are concerned about their market-share and consequently want to cover as many services as possible. However, as financiers, scope and price of the covered package are important for they determine to a high degree the premium to be paid and the insurance company's profitability.

Because many of the three participants are concerned about the effects of an increasingly market-based health care system, in particular with respect to price and scope of coverage, premiums, and contracts that will become a subject of negotiations, each party has started to increase its power base. Hospitals and regional health care insurance companies, for example, have been involved in merger-processes (within their own line of business: mergers between providers and financiers of health care are prohibited in the Netherlands) to increase their size and decrease the number of competitors on the market. Faced with this increase in size and power of the insurance companies, individual providers such as dentists and general practitioners, have become worried that they are in no position to negotiate as individual providers of care with such large insurance companies (Goote, 1990; Scholten, 1990). As a consequence, general practitioners consider the establishment of locally based groups to prevent insurance companies from terminating contracts with individual general practitioners.

Having described the aspects of the Dutch Health Care System that have affected the wear and tear of regional health care insurance companies in the late 80s, account can be taken of what exactly these influences (e.g. the reduction of the number of 'ziekenfondsen' due to the merger-processes that have taken place, and are still taking place) have been on regional health care insurance organizations. The next section therefore, will provide a more detailed description of how 'ziekenfondsen' have become what they are today.

### 5.2.3 REGIONAL HEALTH CARE INSURANCE COMPANIES OR 'ZIEKENFONDSEN'

#### *Historical overview*

The first insurance companies that were established in the Netherlands were so-called 'guild insurance companies'. These guild insurance companies started around 1600 to ensure that employees participating in these guilds, were insured against loss of income during illness, and costs of health care (Boot & Knapen, 1984; Toen en nu. Ziekenfondsen en VNZ, 1984).

In the eighteenth century, however, these guilds were banned by the French (at the time of the French Revolution), and guild-based insurance companies were replaced by commercially-based insurance companies whose objectives were to be profitable. Lack of governmental supervision together with concern for profitability resulted in a situation were both providers and patients became increasingly worried about abuse with respect to the quality and costs of care and insurance. As a consequence, physicians and pharmacists started their own insurance companies, called 'ziekenfondsen' in the middle of the nineteenth century. These 'ziekenfondsen' were non-profit institutions and participation in these 'ziekenfondsen' took place on a voluntary basis. It is important to note that the Dutch federal government did not participate in the establishment of 'ziekenfondsen' at all.

In 1904, however, a so-called 'ziekenfonds' bill was introduced in parliament to settle issues such as the legal status of the physicians and the composition of the ziekenfonds's Board of Governors. However, all bills that were introduced between 1904-1939 to regulate the wear and tear of 'ziekenfondsen', including the above mentioned one, failed to pass the house of parliament. It was only in 1941 (during the German occupation) that a 'ziekenfonds' bill became effective (Boot & Knapen, 1984; Maesen, van der, 1987).

As a result of the 1941 bill, the 'ziekenfonds' insurance, that used to be a voluntary insurance, became a mandatory one. Employees and their families were automatically and willy-nilly insured against the costs of health care. Moreover, the federal government became increasingly involved in the Dutch Health Care System, as reflected in the introduction of a National Council for Health Care [Centrale Raad voor de Volksgezondheid], and the more predominant role government took on in the supervision of these 'ziekenfondsen'. Also, a national 'ziekenfonds' council [Ziekenfondsraad] was installed to manage the central cash in which all the money raised by the premiums for 'ziekenfonds' insurance was brought together. Other activities carried out by this national 'ziekenfonds' council concern the approval of contracts between individual regional 'ziekenfondsen' and their contractors (i.e. the regional providers of health care), and the supervision of the regional 'ziekenfondsen' (Boot & Knapen, 1984).

The effects of the changes introduced by the 1941 bill are first of all that 'ziekenfondsen' are no longer considered as philanthropic activities initiated and ran by benefactors. Secondly, the bill resulted in a substantial improvement of the services covered by the insurance companies both in terms of quantity and quality. Moreover, the 1941 changes led to an improvement of the financial basis of the 'ziekenfonds' insurance system since it had both employees and employers pay for it rather than employees only. And finally, the number of people that fell under the 'ziekenfonds' insurance increased dramatically because the bill introduced a uniform income limit that was substantially higher than its predecessor (Toen en nu. Ziekenfondsen en VNZ, 1984).

In 1966, after a long period of careful preparation, the 1941 'ziekenfonds' bill was adjusted to Dutch legislation and accepted by parliament.

In order to understand why 'ziekenfonds' insurance companies have been able to build good working relationships with the providers of health care, one should realize that 'ziekenfondsen' operate in one particular area. Because of the regional character of their work area, they have been able to closely monitor health care production and consumption, and to visit the providers of care to discuss issues related to quantity and quality of care.

Although most of the work carried out by 'ziekenfondsen' can be considered as administrative in nature (e.g. examining whether or not particular services are covered by the 'ziekenfonds' insurance package; paying the providers of care for their services) they have to some extent been involved in policy making as well (e.g. making agreements with providers of health care regarding the fees paid for their services,[59] giving advice with respect to the planning of hospitals, the number of specialists etc.) (Toen en nu. Ziekenfondsen en VNZ, 1984).

### *Current developments*

As already described in the section concerning the Dutch Health Care System, attempts were made in the second half of the 80s to change the Dutch Health Care System from a government-led system into a market-based system. In this section, account will be given of the major effects these changes (proposed by the so-called Dekker committee) will have (and already have had) on regional health care insurance organizations, in order to arrive at an understanding of the reasons why our client (also a regional health care insurance organization) was willing to participate in the present study.

As previously described, one of the core measures to arrive at market-based health care system, is the introduction of a basic insurance package, which will be of a mandatory nature. However, policy holders will be given the opportunity to select a health care insurance company oneself rather than being compulsory registered at a particular health care insurance company. As a result, the distinction between private and mandatory 'ziekenfonds' insurance will disappear ('ziekenfonds' insurance used to be mandatory) and both private and former 'ziekenfonds' insurers will have to compete for one and the same market. However, since the premium paid for the basic insurance package is set by government (as a percentage of the individual subscriber's income), competition can only take place at the level of additional insurance packages and on the basis of that part of the basic insurance package that is being paid for directly by the policy holders (the so-called nominal premium).

As a result of this change in context (from government-led to market-based), 'ziekenfondsen' have started to realize that a client-orientation will become increasingly

[59] Most of the negotiations concerning the price of services are carried out by the National Council of Regional Health Care Insurance Companies [Vereniging van Nederlandse Ziekenfondsen], representing all individual regional health care insurance companies.

important to maintain, let alone increase, their market-share. This awareness has already led to a more client-friendly approach, diversification of insurance services offered (not restricted to health care insurance alone), and an increase in concern for quality of care (e.g. waiting lists).

Moreover, 'ziekenfondsen' have started to realize that cost-efficiency will have to be taken into account in a system where among others, the price of the product, as reflected in the premium people have to pay, will start playing a role in the selection of an insurance package and the selection of an insurance company.

Because 'ziekenfondsen' will be given the freedom to sign or end contracts with providers of health care, negotiations about the charges for services are expected to become increasingly tough. One way to prepare for these negotiations is to increase one's size as negotiator thereby increasing one's power in the negotiation process. Regional health care insurance companies have responded to the idea of negotiations by merging with other regional health care insurance companies[60].

Another reason why 'ziekenfondsen' have been involved in merger processes lately, is because they have become aware of the economy of scale factor. 'Ziekenfondsen' have come to realize that their performance (in terms of efficiency and client-orientation) is affected by the size they have, for a minimum number of policy holders seems to be required to have say, a marketing and/or automation department of their own[61].

To prepare for all these changes and to decide on what actions to take, it is felt by many 'ziekenfondsen' that their policy making capacity should be improved. Traditionally, 'ziekenfondsen' were mainly concerned with the administrative aspects of health care insurance. These administrative aspects used to be determined to a large extent by governmental regulations, as a result of most policy making activities carried out by regional health care insurance organizations were of a reactive rather than a proactive nature. However, in light of the changes proposed by the Dekker-committee, a more proactive approach seems to be required to survive on a market that will increasingly become a competitive one.

To summarize the changes that have taken place in 'ziekenfonds' land due to the plans proposed in the 'Willingness to Change' report, many 'ziekenfondsen' have started to realize that client-orientation, cost-effectiveness (in relationship to premiums paid by policy holders), and negotiations will become increasingly important to arrive at the low costs and high quality of care that seem required to maintain (or increase) one's market-share. To achieve these objectives, it is felt that an increase in the size of the average 'ziekenfonds' (e.g. by getting involved in a merger-process), an improvement and expansion of the services offered by the organization, and an improvement of the organization itself is required. One of the aspects of such an improvement of the organization itself concerns its policy making capacity, so that an increase can be brought about in the organization's understanding of what is going on and what has to be done to respond to the changing situation in an adequate manner[62].

Now that a description has been given of the changes that are (and have already been) taking place in the context of the organization participating to our participative policy modelling program, account will be given of the participating organization itself to arrive at the exact reasons why it was willing to participate.

---

60 Recently, all sort of alliances have been formed between former regional health care insurance organizations and nationally operating commercial health care insurance companies.

61 Some people say that 'ziekenfondsen' should have at least one billion participants to be able to work in a cost-effective way. Since the Netherlands has 15 billion inhabitants, a dramatic reduction of the number of 'ziekenfondsen is expected for in 1984, the number of 'ziekenfondsen' in the Netherlands was equal to fifty.

62 Recently, our client organization has started to merge the medical and financial-administrative departments in order to be able to provide more integrated services.

## 5.3 PARTICIPATING 'ZIEKENFONDS'-ORGANIZATION

The regional health care insurance organization (ziekenfonds) that was willing to participate in the program can be considered as a relatively progressive 'ziekenfonds'. This progressiveness is reflected in the pace at which the organization had started to prepare for the changes proposed in the 'Willingness to Change' report. A merger process had already been completed[63] and options concerning the issue of market-diversification had already been discussed at the time that we entered the organization. Moreover, the client organization was the first 'ziekenfonds' in the Netherlands that had a research department where research was carried out on issues such as the costs of health care, prescription-patterns, and referral-behaviour. The construction of a conceptual model representing the major factors responsible for the rise in costs of health care has been one of the most striking products of this department so far[64,65]. It was expected that the knowledge incorporated in the conceptual model could be an advantage in a market in which costs would become increasingly important.

In sum, the participating organization can be characterized as an organization that has been aware for some time of the impact many of the changes, proposed by the Dekker committee, can have on the organization itself and the market in which it is currently operating. Hence the organization has been active in looking for ways to prepare for these changes, to increase the likelihood that they will work to their advantage rather than their disadvantage, based on the belief that, if properly prepared, these changes can be considered as opportunities rather than threats.

To understand why participative policy modelling was selected as a tool for this, recall that the research department had been involved in the construction of a conceptual model concerning the costs of health care. It was felt by the research department that this model-building project had been very successful in that it had provided them with a new (and richer) perspective on the problem. By having other departments participate in a model-building procedure as well, it was expected that the knowledge comprised in the conceptual model could be shared with the other departments. One of the reasons why the organization was willing to participate thus was because of the positive experiences it had had with the construction of the conceptual model (they considered the process of building it to be very valuable), and the knowledge that had resulted from it (its product). It was felt that two other departments (the economical-legal [Overeenkomsten & Tarieven] and medical [Medische Dienst] department) should participate in the construction of a model as well to increase their knowledge of the health care system to prepare for their new role of policy maker. For reasons already explained in Chapter Two, it was decided to have the other departments construct their own model as much as possible, rather than having them copy the existing conceptual model, developed by the research department.

The second reason why the organization was willing to take part in the participative policy modelling program is related to the diversity that existed within the department of Care [Zorg]. The department of Care can be considered as a kind of umbrella department, made up of the economical-legal and medical departments. Its diversity not only stems from the fact that it is made up of two departments with different tasks and people (they differ in educational background and experience), but is also related to the fact that it, due to the merger-process just completed, contains people with different

---

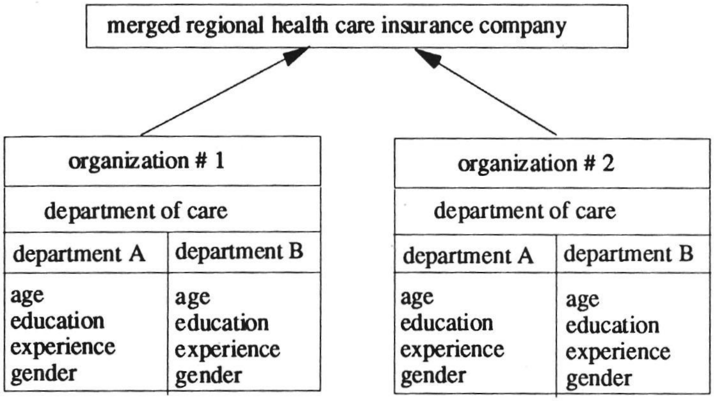
63 Although the merger was completed in 1989, it officially was not until the first of January 1990 that the two organizations had merged into one new health care insurance company.

64 The conceptual model was developed together with researchers from the University of Nijmegen and the University of Utrecht. The model was constructed using a delphi-questionnaire and two structured workshops (Post & Vennix, 1992; Vennix, Gubbels, Post, 1986a, 1986b; Vennix, Gubbels, Post, Poppen, 1988b).

65 The conceptual model was used in the present study as a basis for the preliminary model that was offered to the participants in the first workbook to increase the pace of the participative policy modelling procedure.

organizational backgrounds. The diversity within the department of Care [Zorg] can be depicted as follows:

Figure 5.2: Diversity




To arrive at generally agreed upon (accepted by both departments) policies regarding the department of Care, it was felt that reduction of these differences between the participants and integration of perspectives would be an important step.

Summarizing the reasons why the organization was willing to participate in our program, the program was first of all looked upon as an opportunity to increase one's knowledge about the structure and dynamics of the Dutch Health Care System so that the departments' policy making endeavours would be based on knowledge rather than on intuition only. The participative policy modelling program thus was used to increase the participants' **knowledge** of the complex health care system to prepare for the more active policy making role that the organization was expected to play in the near future. Secondly, the organization aimed to reduce the differences between the two departments that make up the department of 'care', to arrive at a generally agreed upon model of the health care system, a **common perspective** or frame of reference, by means of which issues related to the costs of health care can be discussed. As such, the first reason is related to the notion of individual change, whereas the notion of commonality or homogeneity is being referred to by the second reason.

Now that account has been taken of the reasons why and the context within which our client 'ziekenfonds' decided to participate, an overview will be given of the diversity that existed among the participants at the time the participative policy modelling sessions were taking place. This overview is based on the answers given to some of the questions that were included in the third questionnaire (cf. Appendices 1 to 4).

### *Description of the participants*

To illustrate the diversity that existed among the participants, we will first of all focus on the diversity that existed within the whole group, that is, the diversity that was found among all the people who participated in the participative policy modelling program. Following this description, an overview will be given of the major differences between the medical and economical-legal departments. For this, the very same background variables used to illustrate the diversity within the total group of participants will be employed.

Figure 5.3: Description of the participants: total group (T), Financial-legal department (F), and Medical department (M)

| age | T | F | M |
|---|---|---|---|
| 20-29 | 10 | 10 | |
| 30-39 | 6 | 4 | 2 |
| 40-49 | 3 | 2 | 1 |
| 50-59 | 6 | 1 | 5 |
| >59 | 2 | | 2 |
| total | 27 | 17 | 10 |
| S [5] | .92 | .67 | .76 |

| education | T | F | M |
|---|---|---|---|
| secondary | 4 | 3 | 1 |
| MBO [1] | 4 | 4 | |
| HB [2] | 7 | 5 | 2 |
| university | 12 | 5 | 7 |
| | | | |
| total | 27 | 17 | 10 |
| S | .92 | .99 | .58 |

| years of service | T | F | M |
|---|---|---|---|
| 0-1 | 4 | 3 | 1 |
| 2-3 | 5 | 4 | 1 |
| 4-5 | 1 | 1 | |
| 6-10 | 5 | 4 | 1 |
| >10 | 12 | 5 | 7 |
| total | 27 | 17 | 10 |
| S | .86 | .94 | .58 |

| policy experience [3] | T | F | M |
|---|---|---|---|
| >79 % | 2 | 1 | 1 |
| 60-79 % | 6 | 4 | 2 |
| 40-59 % | 5 | 3 | 2 |
| 20-39 % | 8 | 4 | 4 |
| <20 % | 6 | 5 | 1 |
| total | 27 | 17 | 10 |
| S | .95 | .94 | .91 |

| gender | T | F | M |
|---|---|---|---|
| male | 22 | 12 | 10 |
| female | 5 | 5 | |
| | | | |
| | | | |
| | | | |
| total | 27 | 17 | 10 |
| S | .69 | .87 | 0 |

| specialism [4] | T | F | M |
|---|---|---|---|
| g.p. | 8 | 6 | 2 |
| hospital | 7 | 5 | 2 |
| none | 7 | 5 | 2 |
| sp + hosp. | 3 | 1 | 2 |
| g.p. + hosp. | 1 | | 1 |
| total | 26 | 17 | 9 |
| S | .90 | .78 | .98 |

| organization | T | F | M |
|---|---|---|---|
| # 1 | 18 | 4 | 14 |
| # 2 | 11 | 6 | 5 |
| total | 29 | 10 | 19 |
| S | .99 | .97 | .83 |

1) MBO stands for 'Middelbaar Beroeps Onderwijs, technical and vocational training for16-18 year-olds
2) HBO stands for 'Hoger Beroeps Onderwijs', technical and vocational training for 18+
3) The amount of time they spend on policy making related activities (percentage of their working time)
4) Only the categories that were used are represented in the table. Not mentioned and thus not included were the categories 'sp' and 'g.p. + sp'.
   sp = medical specialist
   hosp = hospital
   g.p. = general practitioner (family doctor)
5) $S = -\sum p_i \log p_i / \ln i$

Figure 5.3, shows that, as far as the total group (T) is concerned, the participants vary in age, educational background, years of service, the amount of time they spend on policy making, gender, and the area of the health care system they consider themselves to be expert in. As such, it is plausible that if these people have to develop policies with respect to, say, the costs of health care, differences in knowledge and opinion do exist. The only variable that does not show too much variety is the variable of gender. Most of the people that work at this particular regional health care insurance company are male. The only females that do work there, work in the economical-legal department, and (note that this cannot be inferred from the information presented in figure 5.3), are of a relatively young age (one of them is falling in the 30-39 category while all the others fall in the 20-29 category).

To illustrate the diversity within the total group and within each of the two subgroups that exists on each of the variables, use can be made of the notion of entropy[66] (Theil, 1970). Note that the relative entropy scores (S) used to express the degree of dispersion of each of the background variables[67], can vary from 0 to 100 per cent, whereby 0 stands for the minimum (all subjects in one category only) and 100 for the maximum of dispersion (all subjects equally distributed over the categories). Since almost all the entropy scores of figure 5.3 are higher than 85 per cent, indicating that 85 per cent of the maximum dispersion is found on each of the background variables, the entropy scores serve to illustrate that indeed heterogeneity was found among the group(s) of participants.

The differences between the two departments obviously are to some extent related to the nature of their activities. The activities that are carried out at the economical-legal department are very much of an economical, administrative and juridical nature, related to the financier's role that regional health care insurance companies play As a consequence, most of the people that work at that department, have an economics or accounting background. Because working for this department does not require a lot of field experience (they need not have been an accountant at a hospital in order to be a good accountant at the regional health care insurance organization), in contrast to the work the people at the medical department do, the average age of the people that work for the economical-legal department tends to be substantially lower than the people that work for the medical department, as shown in figure 5.3. Not surprisingly, the difference in years of service is more or less equivalent to the difference in age for people that are older, usually have more work experience[68].

Regarding the size of the two departments, note that the economical-legal department consists of 17 people, whereas the medical department is made up of 14 people. Of these 14 people, one refused to participate because he felt that the participative policy modelling method was not the best solution to their problem One did participate but was unable to fill out the pre- and posttest because it was almost impossible to get hold of him, due to the nature of his work. Another potential participant could not participate because of absence (holiday). The fourth person who was unable to take part in the program was the person who had already been involved in the construction of the

66 A measure of dispersion for nominal variables with I categories ($i=1, ..,I$) is $-\Sigma i\ p_1 \ln p_i$ The value of this measure is equal to 0 in case the probability of a score is 1 for one particular category, and 0 for all other categories because all subjects fall within one and the same category The maximum value this measure can get is equal to ln I, which is the case if all subjects are equally distributed over the categones However, since the number of categories is not the same for each of the background variables, as a result of which they differ in their maximum value of dispersion (ln I), it was decided to divide the dispersion value by this maximum, to arrive at a relative measure of dispersion, which will be called S

$$S = \frac{\Sigma\, p_i \log p_i}{\ln i}$$

Note that $p_i$ stands for the proportion of subjects in category i and ln for the natural logarithm

67 In order to be able to compare the dispersion scores of each of the variables, it was decided to use one measure of dispersion only. As a consequence, all variables have been regarded as nominal variables.

68 Many of the people that joined the medical department first had their own practice as, say general practitioner, physiotherapist, dentist, or medical specialist. Hence, although the average age of the people of the medical department is substantially higher than the average age at the other department, it need not mean that the people at the medical department have been with the firm for a longer period of time than the people that work for the economical-legal department.

conceptual model in his capacity of head of the research department. As a consequence, the total number of people from the medical department that took part in the program was equal to 10 rather than 14.

The activities that are carried out by the medical department are related to the insurer's roles that regional health insurance companies play. Looking after the quality of health care and dealing with (individual) services offered to their policy holders are some of the activities the medical department is responsible for. Because good working relationships with the providers of care are required to discuss, for instance, the necessity of a service offered to a particular policy holder, it is important that the people who work at the medical department have good medical knowledge (they must have been to a medical school and have taken their MD), and preferably have substantial experience as a practitioner as well. This to ensure that they speak the very same language as their contractors do. As a consequence, many of the people that work for this department previously worked in either their own practice or a group practice.

To sum up, the differences between the two departments are considerable. They are probably mostly related to the difference in the kind of work they do: the medical department being primarily concerned with the insurance company's insurer's role, whereas the economical-legal department is taking care of the organization's financier's role. As depicted in figure 5.3, this results in a difference in age, educational background, years of service, and gender.

## 5.4 A MODEL OF THE DUTCH HEALTH CARE SYSTEM

### 5.4.1 INTRODUCTION

To describe the model of the Dutch Health Care System that was used as a preliminary or 'throw-away' model to speed up the model-building process in the first session, a brief presentation of the study carried out by Vennix et al. (Post and Vennix, 1992; Vennix & Gubbels, 1988a; Vennix & Gubbels, 1992; Vennix, Gubbels and Post, 1986a, 1986b; Vennix et al., 1988b) will be given first. The reason for this is that the conceptual model of the Dutch Health Care System constructed by Vennix et al. has been used heavily in the construction of the present study's preliminary model. Regarding the Vennix et al. model, account will given of both the process by means of which the conceptual model had been constructed, and the product (conceptual model) it has resulted in. Following the description of the conceptual model, the way in which a conceptual model is transformed into a computer model will be examined - attention will be focused upon the conceptualization, formalization, and quantification phases that make up the transformation of the conceptual model into a preliminary computer model. Once the major steps taken in these phases have been outlined, and hence an overview of the construction process has been given, the preliminary model itself, that is, the structure and dynamics of the model, will be presented in some detail.

### 5.4.2 BACKGROUND OF THE MODEL

In 1987, the research department of our client regional health care insurance organization decided to carry out some research on the number of patients referred by the general practitioners to the medical specialists. Attention was being paid to this kind of patients because it was felt that a reduction of the number of patients referred to the more expensive second echelon (the echelon where polyclinical treatment is provided by medical specialists: the outpatients' clinic) would lead to a reduction of the costs of health care. The reason why the number of referred patients does affect the number of patients treated by the medical specialists dramatically, is because patients can only visit a medical specialist or hospital if they have been referred by a general practitioner beforehand (except for emergencies such as a heart-attack). A visit must be paid first to the patient's general practitioner to acquire a referral note saying that one is entitled to be treated by a medical specialist. General practitioners thus act as gatekeepers whose responsibility it is to prevent people from making unjustified use of expensive specialistic treatment.

The research carried out by the research department revealed that considerable differences exist in the way in which general practitioners act as gatekeepers. It was found that some of the general practitioners in the region refer substantially more (a higher percentage of their patients) patients to a medical specialist than their colleagues do. Based on this information, a call was paid to the general practitioners referring substantially more patients to provide them feedback on their referral behaviour in an attempt to lower their referral percentage to lower the costs of health care. However, when confronted with the data on their referral behaviour, many of the general practitioners simply stated that their referral behaviour could not be changed because their population was unlike other populations (e.g. more patients of 65 and older), was located farther away from hospital, or simply contained more patients with a disease that had to be treated in hospital. To avoid having to discuss the impact of factors such as these, the research department first had in mind to have a statistician control for these factors in the calculation of a referral percentage for each of the general practitioners. When discussing the assignment with the statistician, it soon occurred to the research department that they had not considered the fact that medical specialists may react to such a measure (especially if it turns out to be an effective one, that is, one which is able to decrease the number of patients referred by general practitioners) in their discussion of the expected effects of the measure. Medical specialist may, for instance, want to intensify their treatment (order their patients back more often, carry out more medical transactions) in order to make up for this loss in workload and income. As a consequence, it was decided to build a conceptual model of the health care system to acquire some understanding of the way in which the major elements of the health care system are related to each other in order to be able to examine the dynamic effects of measures such as the reduction of the referral percentage, on all three echelons of the health care system - the general practitioner section, medical specialist section, and hospital section - rather than concentrating on the effect on one section only, for it may well be that a reduction of the costs in one section can only be accomplished at the costs of an increase in the costs of another section.

### *The conceptual model*

To arrive at a conceptual model of the Dutch Health Care System, Vennix et al. (1990) used a three-phased approach, in which a preliminary conceptual model was constructed first, followed by a delphi-study, and two structured workshops. Because the information needed to build a conceptual model is often scattered among many different people (experts often know a lot about only a small aspect of the system), it was decided to have a large number of experts participate in the (conceptual) model-building process (Vennix, Gubbels, Post, and Poppen, 1988). As such, the construction of the conceptual model does not differ much from the construction of the computer model of the present study - both stress the importance of participation in eliciting the knowledge required to construct the model. The participation differed however, in that in the present study only people from one and the same organization were invited to take part in the process, while in the former project, people from many different kinds of health care institutions (policy makers, providers etc.) were given the opportunity to participate.

To encourage participation, Vennix et al. decided to construct a preliminary conceptual model of the Dutch Health Care System themselves, on the basis of relevant literature and general insights. This because it was expected that it would be easier for experts to react to an already existing (conceptual) model than to start considering the health care system from scratch.

Once the preliminary conceptual model had been constructed (Vennix, et al., 1986a, 1986b), a delphi-study was conducted, to consult experts and ask them how they felt about the preliminary conceptual model. In this delphi-study, participants were asked to state whether they agreed or disagreed (and why) to a number of statements concerning binary relationships, i.e. relationships between two variables, included in the preliminary model. Based on their answers, a second cycle of the delphi was carried out, in which participants were given a workbook in which they were given the opportunity to comment on a submodel of the Dutch Health Care System rather than commenting on binary relationships only.

Following the delphi-study, a structured workshop was held to enable the participants to interact with each other and discuss their own viewpoints, in particular on issues where the delphi-study had shown a substantial disagreement among the participants.

Based on the information gathered in the above-mentioned phases, a 'final' conceptual model was constructed by Vennix et al. In this conceptual model, three sections were discerned: a general practitioner section (including the people who do not make use of the health care system at all), a medical specialist section (outpatient visits), and an hospital section (inpatient admission). Within each section, three layers were discerned: a layer called 'patients flows' (representing the volume of patients that make use of the health care system), a layer called 'influencing factors' (factors that affect the patients flows and are themselves affected by patients flows), and a layer called 'cost' representing the costs factors included in the model. For a more detailed description of the process by means of which the conceptual model was constructed and the conceptual model itself, the reader is referred to Post & Vennix (1992), Vennix & Gubbels (1988), and Vennix et. al (1986a, 1986b).

The preliminary model presented to the participants in the first workbook, is based on the conceptual model constructed by Vennix et al. The reason why it was decided to pursue the model-building process a bit further, by formalizing and quantifying the conceptual model to arrive at a computer model of health care system, is because a computer model allows one to examine the *dynamic* behaviour of a system, something which can hardly be done with a conceptual model.

Analyzing the dynamic behaviour of the system (including its responses to policy measures) was considered to be of importance by our client regional health care insurance organization because they expected that, in line with the Dekker-report, policy making would become more and more important for regional health care insurance organizations to survive in an increasingly competitive market, as has been explained in detail above. Hence, both the organization's research department and general management were of the opinion that the knowledge acquired during the (conceptual) model-building process should be made accessible for the organization's medical and economical-legal departments to, first of all, increase their *knowledge* of the Dutch Health Care System from a system's point of view, and, secondly, arrive at a kind of a *shared understanding* of the elements and processes taking place in the health care system in which they operate.

Having outlined the major elements of the conceptual model on the basis of which the computer model was built and having explained why it was decided to construct a computer model even though a conceptual model had already been developed by order of the organization's research department, a description will be given of the process by means of which the conceptual model was transformed into a computer model. The description will follow the major phases of the model-building process, that is, the phases of conceptualization, formalization, and quantification. Despite the fact that most of the conceptualization had already been carried out in the construction of the conceptual model, account must be taken of the conceptualization phase as well because, in order to arrive at a manageable computer model, suited to the policy making context in which it was intended to be used, some conceptual adjustments had to be made to reduce the conceptual model's complexity. Following the description of conceptualization, an overview will be given of the formalization of the model. However, since the method used to formalize the model will only be described briefly, the formalization will be discussed in combination with the quantification of the model, that is, the description of the data needed to run the formalized model. In the description of the formalization and quantification, some of the difficulties encountered in our attempt to transform the qualitative model into a quantitative one will be focused upon Note that the phases of sensitivity analysis and validation will not be discussed at this stage, for an overview of the model of the health care system in its totality needs to have been given first It is only when the model of the Dutch Health Care System has been presented that the topics 'sensitivity analysis' and 'validation' can be dealt with

### 5.4.3 CONSTRUCTION OF THE COMPUTER MODEL

#### *Conceptualization*

The first step in the process of transforming the conceptual model into a computer model consists of a reconsideration of the elements included in the conceptual model to determine whether they really need to be incorporated in the computer model There are two reasons why this is important

Firstly, from a policy maker's or participant's point of view, the complexity of the model should be reduced for the participants will have difficulty analyzing and understanding the model's structure (there simply are too many details to pay attention to), let alone that they will be able to grasp the model's dynamic behaviour, if all the elements of the conceptual model are included in the computer model To ensure that the model is used by the participating policy makers, it was felt the model's complexity had to be reduced dramatically

The second reason why the complexity (the number of concepts and relationships included in the model) of the model can and should be reduced, stems from the world of system dynamics From a system dynamics point of view, endogenous elements, that is internally interdependent relationships, play a more important role than the so-called exogenous factors in the replication of the dynamic behaviour of a system As a consequence, quite a few of the variables and relationships of the conceptual model constructed by Vennix et al, can and thus should be excluded from the preliminary computer model because they are exogenous rather than endogenous in nature This is not to say that exogenous variables such as 'ageing population' are not important in the examination of the system's dynamics[69] However, it is to say that most of the dynamic behaviour of a system is brought about by the internal network of interdependent variables and that hence most attention should be paid to these factors in an attempt to account for the dynamic behaviour of the system

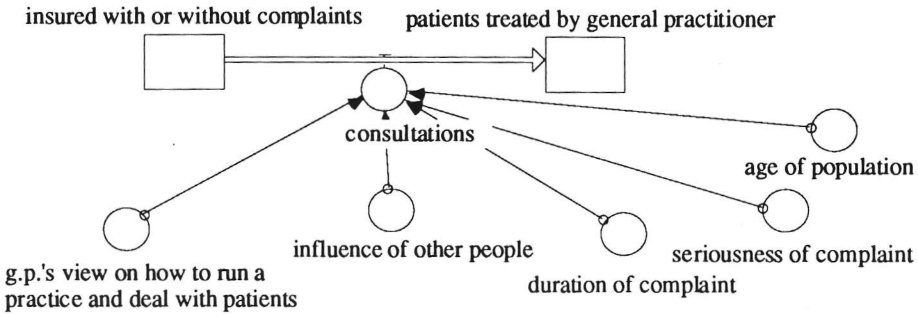
To reduce the complexity of the model developed by Vennix et al, the guideline concerning the notion of endogeneity was applied first To illustrate the effects of the notion of endogeneity on the complexity of the model, look at the differences between the way in which the processes that affect the number of people that pay their general practitioner a visit for a particular complaint have been represented in the conceptual model and the preliminary (computer) model In figure 5 4, an overview is given of the conceptual model of the Vennix et al (1988) study Note that the consultations (and visits[70]) that are focused upon in that model concern the so-called first consults, that is, consultations in which patients see their general practitioner for the first time for a particular complaint

---

[69] The impact of external influences or exogenous factors on the dynamic behaviour of the system was examined in great detail in the second session of the participative policy modelling program

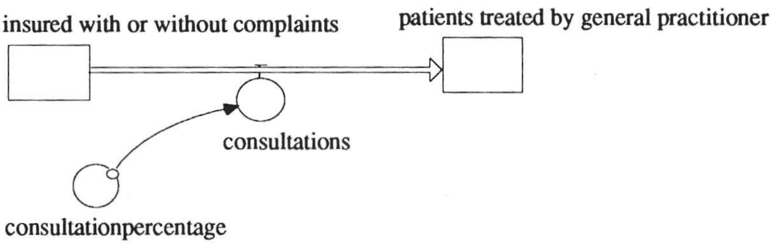
[70] Consultations concern patients visiting a general practitioner whereas visits concern general practitioners paying patients a visit at their homes

Figure 5.4: Conceptual model




Since the variables 'seriousness of the patient's complaint', 'duration of the complaint', 'influence of other people', 'age of population' and 'general practitioner's view on how to run a practice and deal with patients' can all be regarded as exogenous factors (they affect other variables, but are themselves not affected by any of the variables of the model), it was decided to remove all five factors for the time being and replace them by one general variable called 'percentage of people consulting a general practitioner', representing the percentage of people from the level 'insured with/without complaints' that visit the general practitioner for the first time in one particular week:

Figure 5.5: Simplified model




The same procedure has been carried out with respect to many other exogenous variables of the conceptual model developed by Vennix et al.

As a corollary to that, attention has been paid to the feedbackloops that were included in the conceptual model, to ensure that they were incorporated in the computer model, for these circular cause-effect relationships are a key element of system dynamics representations of a system. An example of such a feedbackloop in the conceptual model is the loop between the workload of the general practitioner and number of patients referred back for another visit, say, three weeks later.

Figure 5.6: Feedbackloop in the conceptual model

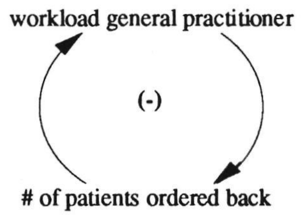
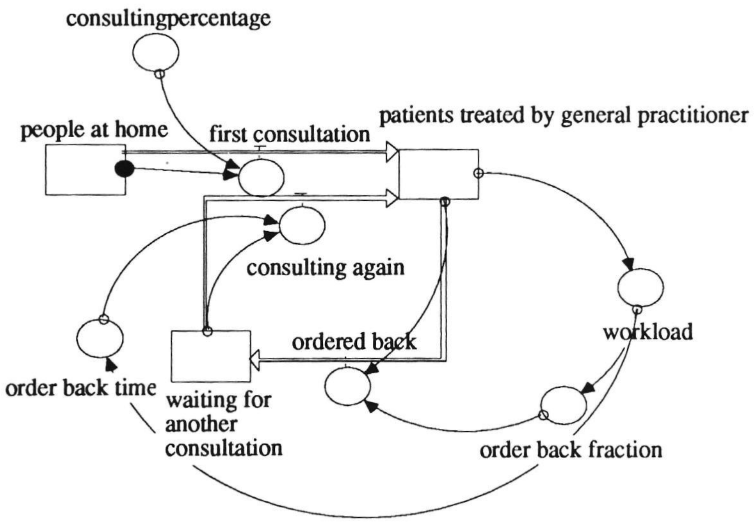



Figure 5.6 shows that an increase in the number of patients ordered back for a second (or third or fourth or...) visit for the very same complaint will lead to an increase in the workload of the general practitioner, for the people who are ordered back, eventually show up at the general practitioner's consulting-hours[71]. However, the more patients show up, that is, the more busy a general practitioner is, the less patients (s)he will order back, for general practitioners are very much aware of the fact that those who are ordered back will come back. To illustrate that the reconsideration of the conceptual model in the conceptualization phase not always did result in a reduction of the complexity, an overview is given of the way in which the feedbackloop depicted in figure 5.6 has been represented in the preliminary computer model:

Figure 5.7: Incorporation of a feedbackloop in the computer model




71 Note that the model does not include a leakage flow representing the patients that decide themselves not to return to the general practitioner, despite the fact that they have been ordered back. Since the number of people who do not return is very small compared to the number of people that do come back, leakage flows like these have not been included in the model for the sake of simplicity.

In figure 5.7, it is shown that general practitioners have two mechanisms to regulate the number of patients paying them a visit rather than the single one presented in figure 5.6.

In the first place, general practitioners have the opportunity to reduce their workload by lowering the number of patients they order back (in the computer model an order back fraction is used for this, saying, that one changes his or her order back behaviour, that is, changes the relative number (percentage) of people that are ordered back). The variable 'people waiting for a second (or third or ...) consultation' is added to the model in order to be able to distinguish between people who visit the general practitioner for the first time, and those who visit due to being ordered back. Those who have been ordered back spend some time in the level 'waiting for another consult, and once they have waited long enough, they flow back to the level of patients treated by the general practitioner - they pay the general practitioner another visit.

The second way in which general practitioners can, and actually do, change their workload, is by varying the so-called order-back time, that is, the time in between two consecutive visits of the patients they order back. For instance, in case an increase in workload has taken place, general practitioners ask their patients to come back for another visit in three rather than the usual two weeks. By doing this, they increase the time that people spend in the level 'waiting for another visit', and (only temporarily though) reduce the number of patients that actually do consult them for a second (or third or...) time[72].

In addition to the notions of endogeneity and circular cause-effect relationships, another guideline of system dynamics tradition was followed to determine which of the elements of the conceptual model had to be included in the computer model. Recall that it was recommended by system dynamicists that a particular problem should be modelled rather than just a system, for focusing on a problem serves to determine what elements to include and exclude from the model. This guideline was followed in a broad sense in that not only the costs of health care was taken into account in the decision what to include and exclude from the model, but also account was given to the purpose for which the model was being constructed (recall that it was intended to be used by policy makers to increase their understanding of the dynamic behaviour of the health care system with respect to the costs of health care).

Keeping in mind the purpose for which the model was constructed, for instance, lead to the decision to remove the distinction in the conceptual model between insured people without complaints and insured people with complaints. Rather than having two levels in the model to refer to these components of the health care system, it was decided to take the two levels together and call them 'people at home', representing the people who do not take part in the health care system to have a particular complaint seen to. Moreover, since waiting time is not an issue for patients visiting a general practitioner (people simply go to the consulting-hours or make an appointment for which there is hardly a waiting time) *and* prevention was not an important issue at the time the model was constructed. The information added by the inclusion of a level in between the people who are at home and the patients that are treated by a general practitioner regarding the number of people waiting and the average waiting-time is not relevant and hence does not offset the increase in complexity it is bringing about.

The levels of patients waiting for a treatment by the medical specialist (polyclinical treatment or outpatient treatment) and patients waiting for admission at a hospital (for

72 The third way in which general practitioners can adapt to a change in number of patients that visit them on their consulting-hours, is by changing the length of a consultation. They, for instance, can try to spend eight rather than the usual ten minutes per patient to be able to see more patients in the same amount of time. This has been built into the model only indirectly, by having the general practitioners not change their order back time and order back fraction immediately, but wait until the change in workload can no longer be dealt with by shortening the average time spent per patient In the description of the content of the preliminary model, an illustration will be given of the way in which this indirect reduction of the workload has been incorporated into the computer model

clinical treatment) by contrast, were included in the computer model because waiting lists and waiting time regarding the medical specialist and hospital sections were topical affairs at the time of the construction of the model[73].

Also, to reduce the complexity of the model, an attempt was made to remove the conceptual model's distinction between patients under treatment for a short period of time (that is for one consultation only), and patients treated for a longer period of time (those who are ordered back). It was felt that a considerable reduction of complexity could be arrived if the two levels were taken together, as occasionally had been done in the presentation of the conceptual model (Post & Vennix, 1992; Vennix & Gubbels, 1988a; Vennix & Gubbels, 1992).

However, it was soon found out that taking the two levels together would lead to an increase rather than a decrease in the complexity of the computer model. Firstly, because of the increase in technical complexity of the model that is brought about if knowledge concerning the number of patients, number of patients ordered back, order back fraction, order back time, and the number of first consults, is to be taken together to arrive at a satisfactory description of the dynamic behaviour of one level. This increase in technical complexity is to some extent due to the fact that in the computer model knowledge has to be represented in terms of continuous rather than discrete flow processes[74].

Moreover, taking into account the people and purpose for which the model was constructed, it was concluded that because the concepts used to describe the two-level solution (e.g. the concepts 'order back fraction' and 'order back time') appeared to be much more in line with the concepts used by the policy makers themselves to refer to the processes taking place, a two-level representation was to be preferred to a one-level solution. However, note that the two levels have been named somewhat differently, primarily to stress the waiting lists that are part of the system.

### *Formalization and quantification*

The formalization and quantification of the conceptual model was started with the patients flows, the flows of patients going from one part of the health care system to another. Not only because these patients flows form the structural core of the health care system, but also because they are relatively well known and had already been presented in terms of levels and rates in the original hybrid model constructed by Vennix et al.

It was followed by the formalization and quantification of the influencing factors, the factors that affect the patients flows and are themselves (most of them) affected by these patients flows as well. An example of an influencing factor is the workload of a medical specialist. It is assumed that the workload is affected by patients flows (e.g. the more patients consult a medical specialist, the higher the workload will be), and that workload can affect the flows of patients as well, for instance, because a medical specialist may decide to alter his or her order back behaviour to compensate for changes in workload.

Once the patients flows and influencing factors had been formalized and quantified, attention was paid to the third layer of the model, the costs produced by the system's patients flows and influencing factors. Recall that the costs were included in the model because one of the major reasons to have a computer model built was to acquire knowledge that could be used to contain the costs of health care. Anticipating the description of the content of the model, note that most of the costs are directly related to the production of care generated by general practitioners, medical specialists, and hospitals. As a consequence, many of the costs follow the patterns of the patients flows if

---

73 At the time the computer model was constructed, research was carried out by order of the National Medical Specialists Association (LSV, 1990) to examine the average waiting time for each of the specialisms.

74 The Ithink modelling package had not been on the market at the time of the construction of our computer model. In contrast to the most well-known packages that were available at that time (STELLA and DYNAMO), Ithink allows one to distinguish between continuous and discrete flows.

changes in price are not taken into account. To illustrate, the costs of visiting a medical specialist, for example, can be represented as the multiplication of the average price of a referral note and the number of patients that flow from the general practitioner to the medical specialist. However, not all the cost factors included in the model are so directly related to the patients flows. The costs of drugs prescribed by a general practitioner, for instance, is calculated by multiplying the number of patients that visit a general practitioner by the amount of drugs prescribed per patient visiting the general practitioner and the average price of the drug prescribed by the general practitioner[75]. Because of the influence of the average amount of drugs prescribed per patient visiting the general practitioner may have on the total number of drugs prescribed, the number of drugs prescribed need not follow exactly the 'number of patients that visit a general practitioner' belonging to the model's patients flows.

### *Formalization*

To formalize the three layers of the conceptual model (patients flows, influencing factors, and costs), use was made of the format offered by the DYNAMO simulation package. The model was constructed in DYNAMO because at that time, both client and researchers were predominantly working on DOS-machines. However, a STELLA version was developed for demonstration purposes as well[76].

In the DYNAMO programming language, explicit reference to time-indices is being made. To illustrate the structure of the DYNAMO language, an example is given of the way in which three of the most common types of formal equations are expressed in it:

Figure 5.8: Three basic equations: level, rate, and auxiliary

| | |
|---|---|
| $L.K = L.J + (R_{IN}.JK - R_{UIT}.JK)$ | LEVEL |
| $R.KL = f(L.K, A.K)$ | RATE |
| $A.K = f(L.K, R.JK)$ | AUXILIARY |

The time-indices are used to refer to the past (J), the present (K), and the future (L). The development of the value of a level, for instance, depends on the value it had (that is, at time J), and the net result of the in- and outflow of the period in between the last value (time J) and the value to be calculated (time K): JK. Note that stocks (levels) thus can be regarded as integrals of flows over time, whereas flows can be considered as time derivatives of these levels[77].

---

75 Some calculations had to be carried out to arrive at data such as the average amount of drugs prescribed by a general practitioner per consulting patient, for most of the data found in literature and research reports focus on the average amount of medicine prescribed per person per year. This to illustrate that not always an exact match existed between the kind of data required to quantify the model and the kind (or level) of data used in literature and research reports.

76 The presentation of the model in the present chapter is based on an IThink version of the model, because it was felt that the IThink package allows for a nicer and easier presentation of the dynamic behaviour of the model (e.g. the effects of sensitivity analyses, the effects of exogenous variables on the steady-state model, and the dynamic effects of policy measures). This is not to say that graphical presentation of the dynamic behaviour cannot be carried out in other modelling packages.

77 For a more detailed description of the relationship between differential equations and the equations used in packages such as DYNAMO, Stella, and Ithink, the reader is referred to Richardson & Pugh, 1985, and Chapter 19 of the IThink User's Guide.

Now that a brief description has been given of the structure of the formalization language used in the formalization of the conceptual model to represent the causal relationships as mathematical equations so that simulations with the model can be carried out[78], account must be given of the quantification of the model, that is, the values attributed to the model's parameters, for it is only when a formal model is quantified that it can be used for simulation purposes.

## *Quantification*

To estimate the value of the parameters of the model, use can be made of three different approaches (Richardson & Pugh, 1981). One can use data on individual relationships in the model, firsthand knowledge of the process, and data from the overall behaviour of the system.

The easiest ways to quantify a model, is to use existing knowledge on *individual relationships or variables* of the model, for they can be applied to the model directly without having to be converted. An example of such knowledge is the knowledge we have about the percentage of patients ordered back by a general practitioner because of the research carried out by Post (1984b).

Using firsthand knowledge of the *process* means that the value of a parameter is estimated 'below the level of aggregation'. Statistics regarding the factors that make up the aggregate model variable are combined to arrive at a value of the aggregated variable. To illustrate how this works, look at the way in which a value of the variable 'average number of drugs prescribed per contact/consultation' can be arrived at:

Figure 5.9: Estimating below the level of aggregation

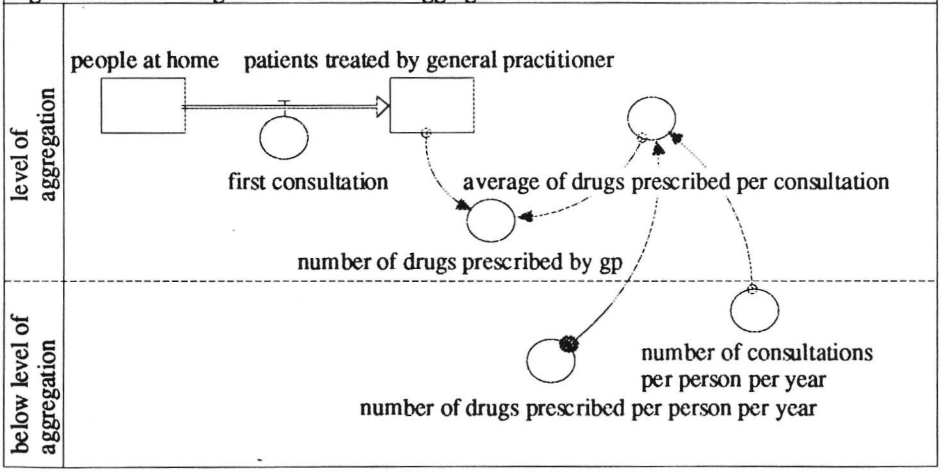


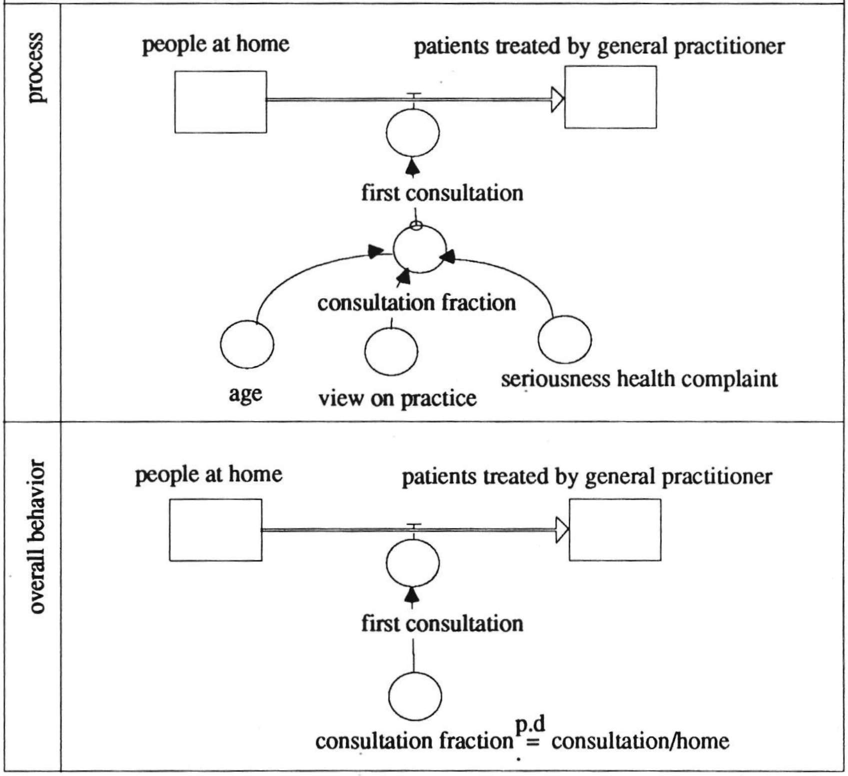

Figure 5.9 shows that a value of the aggregated variable (the number of prescribed drugs per consulting patient rather than per person per year) can be arrived at by combining the values of the variables 'number of prescribed drugs per person per year' and 'number of consultations per person per year'.

The difference between estimating the value of a parameter on the basis of knowledge of the process and knowledge of the *overall behaviour* of the system is that the latter is based on knowledge of the same level of aggregation as the model itself is, whereas the former, as explained before, is based on knowledge of variables that are of a somewhat

78 An overview of the way in which the model has been formalized using the IThink modelling language is given at Appendix 6. The DYNAMO and IThink notations do not differ substantially.

lower level of aggregation. To illustrate, to arrive at a value for the consultation fraction, indicating how many persons of the 'people at home' level consult a general practitioner in a particular period of time, one can either try to find statistics on all the elements that contribute to it (e.g. age, general practitioner's view on how to run a practice and deal with patients), or simply use the knowledge one has about the overall behaviour of that part of the model. In case one decides to use the latter approach, a value of the consultation fraction can be arrived at by dividing the number of patients that are treated by the general practitioner (minus the people who consult for a second or third time) by the number of people who do not take part in the health care system, the people at home. The difference between the two approaches can be depicted as follows:

Figure 5.10: Two approaches to the consultationfraction




All three ways of estimating the value of a parameter have been employed in the construction of the preliminary computer model of the Dutch Health Care System.

Knowledge about individual variables or relationships for instance, was used to arrive at values for variables such as the number of general practitioners, medical specialists and hospital beds[79]. Moreover, direct measures or data were also used for variables such as the fraction of people ordered back, referred, and discharged by general practitioners

79 To arrive at the number of hospital beds for the period between 1983 and 1988, it was expected that it would suffice to simply add the number of beds of each of the hospitals per year, as stated in the hospitals' annual reports. However, some of the data were in conflict, and it was only then that we found out that there were differences in the definition of hospital beds due to the question as to whether empty beds (beds on wards that no longer were used) should be included or not.

(Post, 1984b), and the number of patients in the level 'hospital' (based on the annual reports of five hospitals of the region).

Examples of parameters that were estimated below the level of aggregation are the number of medical transactions carried out by medical specialist, for data on individual kinds of medical transactions were used to arrive at a value for the total number of medical transactions (the kind of medical transactions that were included in this grand total cover 80 per cent of the medical transactions carried out in the Dutch Health Care System. Note that this value, regarding the overall behaviour of the system, in turn, was used to arrive at a value of the parameter 'average number of medical transactions per patient'. Other examples of variables of which the value was estimated below the level of aggregation are 'the fraction of the average number of prescribed drugs per contact' (cf. figure 5.9), and 'the average price of a referral note', for they both are made up of factors that are of a lower level of aggregation.

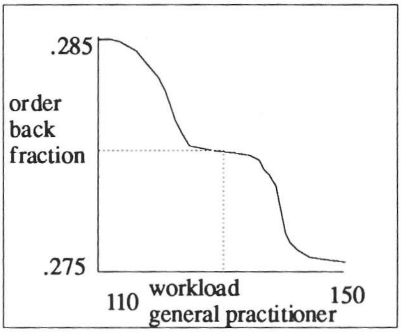
Knowledge of the overall behaviour of the system was used to arrive at values for variables such as 'the number of medical transactions per contact', 'the price of transactions' (being equal to quotient of total costs of transaction and total number of medical transactions), and values for the levels 'people at home', 'people waiting for the general practitioner', 'waiting for the medical specialist', and 'waiting for admission into hospital'. Since no information was available about the length of the waiting list (that is, the number of people waiting), the number of people flowing into these waiting levels was multiplied by the average time people spend in the level, to arrive at initial values for those 'waiting'-levels.

As mentioned before, one of the problems encountered in the quantification of the formal model was that occasionally conflicting values were found (e.g. regarding the number of hospital beds). Moreover, because regional data were not always available, use had to be made of national (or provincial) data every now and then despite of the fact that substantial regional differences in the health care consumption exist (Hoefnagels, 1989; Stichting KLOZ informatiesysteem gezondheidszorg KISG86, 1986, Post, 1990; Speld, van der, 1990). To assess the value of the ratio of visits to consultations at the general practitioner section, for example, national statistics were employed (Vademecum Gezondheidsstatistiek Nederland, 1989, p. 148), and to assess the effects of an ageing population on some of the variables of the model, for instance, use was made of national data represented in reports such as "De Intramurale gezondheidszorg in cijfers per 1 januari" (1986, 1988), KISG86 Jaarboek 1986 (1986), and LMR Jaarboek (1985, 1987). The degree to which the national values hold for the region in which our client health care insurance organization is operating remains to be seen.

Moreover, quantification of the relationships between workload (or occupation rate) and fractions such as order back fraction, discharge fraction, and refer fraction proved to be extremely difficult for no statistics were available on the basis of which these relationships could be quantified. However, as far as the direction of the relationships between workload and the above mentioned fractions are concerned, agreement existed among the people taking part in the Vennix et al. (1988) study. To illustrate, most of the people participating to this study were of the opinion that for instance an increase in the workload would lead to a decrease rather than an increase in the order back fraction. The quantification of this relationship hence is based on a mental experiment carried out on the basis of the agreement mentioned above. The quantified relationship that resulted from this combination of mental experimentation and expert opinion, can be depicted as follows[80]:

80 The relationships of the model are defined and quantified using a table-function. In appendices 6 and 7, an overview is given of all table functions included in the model.

Figure 5.11: Table function workload-order back fraction




As shown in figure 5.11, it is assumed that at some stage, when the experienced workload differs too much from the usual (or desired) workload, general practitioners adjust the percentage of patients they order back. It is assumed that general practitioners do not change their order back behaviour immediately, for they first reduce the average time spent per patient. However, since there are limits to this reduction in time spent per patient, they then start to change, among others, their order back behaviour. Also depicted in figure 5.11 is that there exists an upper- and a lower-limit to the way in which general practitioners can vary their order back behaviour. Obviously, some people have to be ordered back even when the general practitioner is extremely busy, and vice versa, some people simply cannot be order back even when there is plenty of time to see them for, say a second or a third time.

Since exact statistics regarding the relationship between workload and order back fraction are lacking, the relationship depicted in figure 5.11 needs to be regarded as an assumption about the causal relationship between the two variables. To check whether these assumptions make sense, our participants were asked to comment on each of the table-functions included in the model, and to change them if they felt they were incorrect.

Now that an overview has been given of the quantification of the model, account will be given to the model of the Dutch Health Care System itself. An overview will be given of the elements that have been included in the model (i.e. the structure of the model), and they way in which they interact with each other (the dynamic behaviour the structure is able to bring about).

## 5.5 THE PRELIMINARY MODEL

### 5.5.1 INTRODUCTION

In this section, an overview will be given of the preliminary system dynamics model of the Dutch health care system. The description of the model will start with the section of the general practitioner. It will include the patients flows that surround the general practitioner, the influencing factors that affect these flows, and the costs generated by that part of the system. Following the description of the general practitioner subsystem, an overview will be given of the medical specialist subsystem (the part of the system where patients visit a medical specialist at the hospital's outpatient department), consisting of patients flows, influencing factors, and factors concerning the costs of this section as well. Finally, once the patients flows, influencing factors and costs related to the section of the hospital have been dealt with, an overview will be given of the model of the Dutch Health Care System in its totality. In this overview, the basic elements of the model will be depicted and a description will be given of the dynamic behaviour of the model. For this, account will be given of the steady-state model that was constructed, a sensitivity analysis that was carried out with the model, the effects of adding some exogenous factors to the basic model, and the dynamic consequences of a particular policy measure.

Once the dynamic behaviour of the model has been focused upon, the issue of validation will be looked into. The model's validity will be discussed in relationship to the issue raised in Chapter Two concerning the stability of the real world system and the chances of making an error of the third kind (solving the wrong problem).

### 5.5.2 GENERAL PRACTITIONER

#### *Patients flows*

To describe the processes that take place in the general practitioner subsystem of the Dutch Health Care System, an overview will be given of the patients flows first. It is important to realize that the model will not be concerned with individual patients or health care providers. By contrast, the model aims to represent all patients and providers of health care of a particular region, in order to be able to describe the processes that take place on a regional level.

In the Dutch Health Care System, patients are submitted to the following health care processes when consulting a general practitioner:

* The process of consulting a general practitioner usually starts with a health complaint for it is on the basis of a complaint that patients-to-be decide to visit (consult) a general practitioner and become patients in a true sense.

* To visit a general practitioner, one either makes an appointment, or goes to the consulting-hour for which no appointment needs to be made.

* Being treated by a general practitioner means that a diagnosis is being made first, on the basis of which a general practitioner decides whether or not medical transactions need to be carried out (only small medical transactions are carried out by general practitioners themselves) and/or drugs have to be prescribed, or that a piece of advice is given to the patient.

* At the end of a consultation, general practitioners either ask a patient to return for another consultation (order a patient back), referred him or her to a medical specialist, or discharge the person[81].

* Patients who are discharged return to their homes. They are no longer considered as a patient.

* Patients who are referred to the medical specialist, however, will have to make an appointment to see the medical specialist. On average, this takes about four to five weeks. In the presentation of the processes surrounding the medical specialist, a more detailed description will be given of the processes that take place when the patient has waited for the four to five weeks that are required and enters the consulting-hour of the medical specialist.

* Patients who are ordered back do not really leave the health care system, for they are still under treatment rather than being discharged. In terms of the model, this means

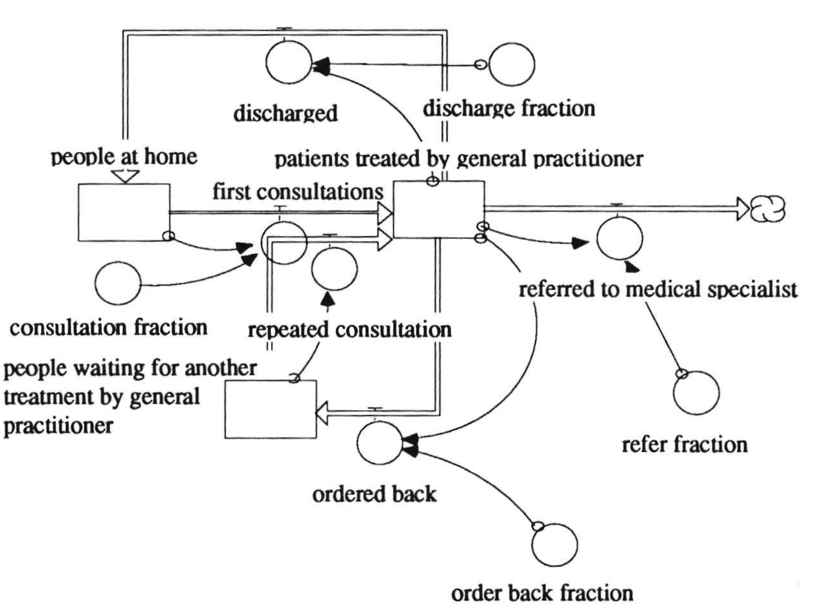
---

[81] Note that patients can also be referred to other health care workers. This flow is not included in the model of the health care system for it was decided to include only the patients flows that are most important in terms of their volume.

that they will spend some time in a level representing the people who are waiting for another treatment by the general practitioner.

The patients flows corresponding to the processes just outlined, have been represented in the preliminary computer model as follows:

Figure 5.12: General practitioner section of the model




Because general practitioners are not paid per treated patient, but per person registered at their practice per year, hardly any regional data existed with respect to the number of patients treated by general practitioners, and the size of health care consumption generated in that section of the system. Although some national statistics are available on privately insured patients (patients whose income is higher than 50.000 Dutch guilders a year), it was decided to have some research carried out rather than employing national data and run the risk of arriving at statistics that do not really apply to the region of our client.

Based on this research (Ree, van der, Mokking, Post, and Gubbels, 1991), a consultation fraction of 0.046 was found, meaning that 4.6 per cent of the people who are not taking part in the health care system (the people in the level 'people at home'), visit their general practitioner in a particular week[82]. Regarding the other three fractions included in the section of the general practitioner, the research carried out by Van der Ree et al. revealed that 13 per cent of the patients treated by a general practitioner is referred to a medical specialist, 28 per cent of them is ordered back, and 59 per cent of the patients who visit a general practitioner in a particular week, is discharged.

Note that the patients who are requested to return for another visit (the outflow called 'ordered back') flow into a level called 'patients waiting for another treatment' in which

82 Note that all values of the model have been standardized to a week. The reason why this was done, is because the model's simulation time is equal to one week. Thus the number of patients treated by a general practitioner, for example, is not considered per day, but per week.

they, on the average, spend 4 weeks before they are allowed to consult their general practitioner again. Since they are still under treatment for the very same health complaint and, strictly speaking, do not leave the health care system, it was decided to represent them as a separate level, rather than having them return in the level 'people at home'.

To understand the dynamic behaviour of the system presented later on, one should realize at this stage that because the size of the rates is the result of a multiplication of the content (or size) of the level and the corresponding fraction, an increase in the size of a flow can be brought about by either an increase in the size of the level or an increase in the corresponding fraction. As a consequence, an increase in, say the level, does not necessarily result in an increase in the flow, for it may well be that this increase in the size of the level is compensated for by a decrease in the corresponding fraction.

The next step in the description of the patients flows surrounding the general practitioner concerns the patients who have been referred back from the medical specialist section or the hospital section to the general practitioner for another consult. These patients enter the level 'patients waiting for the general practitioner', before they visit the general practitioner. Note that because they already had consulted their general practitioner for the very same health complaint (otherwise they could not have consulted a medical specialist), their consultation is not a first consultation but a so-called repeated consultation.

## *Influencing factors*

Now that the most important patients flows of the general practitioner section have been described, account must be taken of the factors affecting these patients flows that have been built into the model.

The most important influencing factor included in the computer model, is the general practitioner's workload. It is assumed that changes in the patients flows of the general practitioner section, if not brought about by changes in say the second echelon, or a change in the size of the total population, are brought about by the way in which the flows of patients who are ordered back, referred, or discharged are related to each other - by changing the fractions of patients that are ordered back, referred, and discharged, general practitioners can change the flows of patients that surround them.

Workload is defined as the quotient of the total number of patients treated by the general practitioners of the region and the total number of general practitioners. By dividing the total number of patients treated per week by the total number of practitioners, an average workload of 130 patients per general practitioner per week is obtained.

If the workload decreases, it is assumed that, at some stage, general practitioners change the ratio's between the order back, refer, and discharge flows to compensate for this change in workload. Since the sum of the three fractions always needs to be equal to 1 (there are no other ways to leave the level except for being referred, ordered back, or discharged), it was decided to have the workload affect the order back and refer fraction directly, and the discharge fraction indirectly.

The basic structure of the negative feedbackloop between workload and patients flows is the following: if a general practitioner considers his or her workload[83] to be too ***high***[84], an attempt will be made to reduce the number of patients visiting the consulting-hours. To do so, it is suggested by the model that use is made of the opportunity to order back less patients, for patients who are ordered back eventually return and thus contribute

[83] In the model used during the sessions, a distinction was made between real and perceived workload. This was done by smoothing the value of the real value of the workload. However, in order to not to complicate the presentation of the model too much, it was decided not to include this distinction in the presentation of the model in the present chapter.

[84] Note that more or less the same line of argument (but in reverse) can be used to explain how general practitioners compensate for a workload which is considered to be too low.

to the (too) high workload. However, the question is what will happen to the patients who are no longer ordered back now that the order back behaviour has been changed (i.e. now that the order back fraction has been lowered). The patients still have to leave the consulting-hour, that is, leave the level 'patients treated by the general practitioner'. It is expected that to allow for such a reduction in order back fraction, substitution will have to take place.

Some of the patients that really have to be dealt with (i.e. patients who absolutely cannot be discharged) can be sent to the medical specialist for treatment, rather than being treated by the general practitioner him- or herself. This will save quite some time. However, only patients who really need to be dealt with by the medical specialist can be referred to the medical specialist if the general practitioner does not want to spoil his or her relationship with the medical specialist. Consequently, the refer fraction can only be increased slightly.

To decrease one's order back rate, general practitioners can also substitute between the order back flow and the flow of patients being discharged. It is common knowledge that many of the people who are ordered back by a general practitioner, need not be ordered back at all. A small percentage of the people who visit general practitioners, account for a large percentage of the visits being paid to general practitioners. Substitution between the order back and discharge flows thus means that some of the people who do not really have to be ordered back are sent home rather than being given the opportunity to return for another visit. As such, a reduction of the order back fraction can be realized at the cost of an increase of the discharge fraction.

As depicted in figure 5.7, general practitioners can also reduce their workload (though only temporarily), by increasing the order back time, that is, ordering their patients back in, say, four rather than three weeks. The reduction is only temporarily because the same amount of people has to be seen to by the general practitioner.

Another factor belonging to the layer of influencing factors is the number of drugs prescribed by practitioners. Although it is considered to be part of the layer of influencing factors, it is only affected by the patients flcws and does not affect the patients flows itself.

The number of drugs prescribed by a general practitioner is the product of the number of consulting patients (s)he is having and the average number of drugs prescribed per contact/consult. According to the model, general practitioners increase the number of drugs they prescribe per consulting patient if their workload increases, and vice versa, lower the number of drugs they prescribe per consult, if their workload decreases. The rationale behind this is that prescriptions can be used to end a consultation without being rude and without giving the patient the feeling that his or her health complaint has not been taken seriously (for a description of the exact relationship between workload and average number of drugs prescribed per consult, the reader is referred to Appendix 7).

Note that the number of medical transactions carried out by the general practitioner has not been taken into account in the model because in the present day's insurance system, regional health care insurance organizations do not pay for these transactions separately. General practitioners receive a fixed fee for the people who are insured at a regional health insurance company, covering all consultations and medical transactions on an annual basis. In 1983, the fee paid to general practitioners was equal to about 100 Dutch guilders per person per year[85].

Summarizing, in the preliminary model, a relationship is found between the workload of a general practitioner and his or her order back fraction, refer fraction, discharge fraction, order back time, and average number of drugs prescribed per consulting patient. The exact nature of those relationships is depicted in Appendices 6 and 7, where an

85 Today; 1 US$ is equivalent to 1.8 Dutch guilders; 1 Dutch guilder costs about .56 US$. Hence, general practitioners get 55 US$ per mandatorily insured patient per year (note that the costs of medicine is not included in this figure).

overview is given of the relationships that have been formalized by means of a table-function. Building these relationships into the model depicted in figure 5.12 results to the following presentation:

Figure 5.13: The general practitioner section of the model

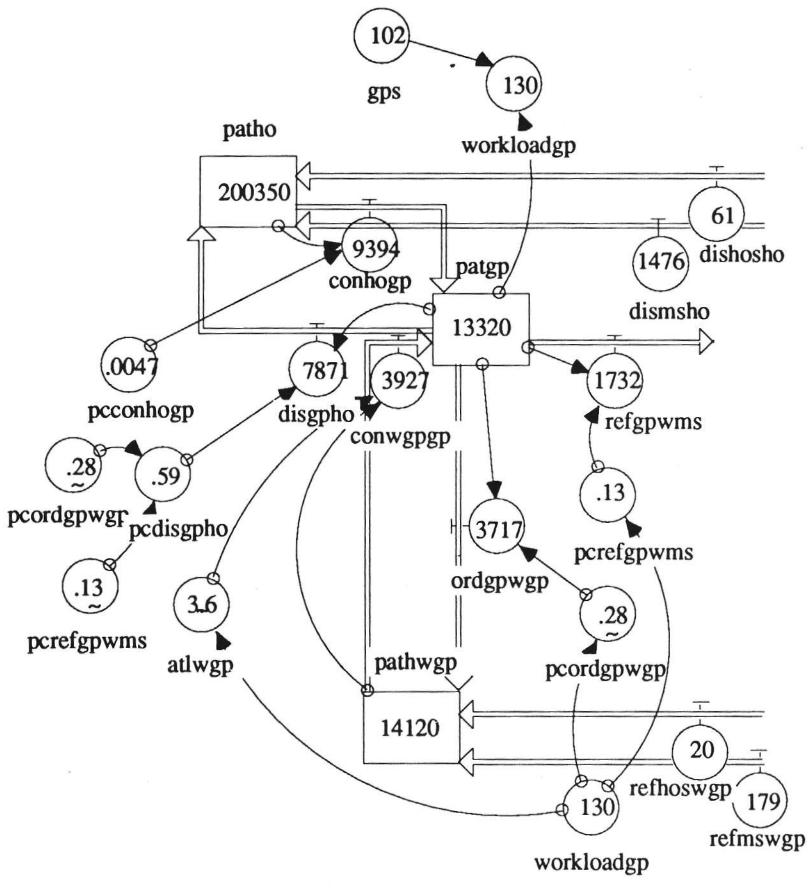


Figure 5.13 (cont.): abbriviations used in the figure on the left

Levels:
patho = people at home
pathwgp = patients waiting for another treatment by the g.p.
patgp = patients visiting a general practitioner

Rates:
disgpho = discharge (from level g.p. to level people at home)
refgpwms = refer (from level g.p. to level waiting for medical specialist)
ordgpwgp = order back (from level g.p. to level waiting for g.p.)
refhoswgp = refer back (from level hospital to level waiting for g.p.)
refmswgp = refer back (from level medical specialist to level waiting for g.p.)
dishosho = discharge (from level hospital to level people at home)
dismsho = discharge (from level medical specialist to level people at home)
conhogp = consultation (from level home to level g.p.)
conwgpgp = consultation (from level waiting for g.p. to level g.p.)

Auxiliaries:
gps = number of g.ps
workloadgp = workload of g.p.
pcrefgpwms = percentage of patients referred from g.p. to medical specialist
pcordgpwgp = percentage of patients ordered back by g.p. to level waiting for g.p.
pcconhogp = percentage of people consulting a g.p.
pcdisgpho = percentage of people discharged by g.p.
atlwgp = average waiting time in level patients waiting for g.p.

The initial values of the variables have been included as well, to provide an overview of the quantities of this section of the model. Note that the quantities have been rounded of slightly for presentational purposes. For an overview of the exact values of the parameters of the model, the reader is referred to Appendices 6 and 7[86].

## *Costs*

The costs generated in the general practitioner section of the health care system consist of:

* costs of consulting a general practitioner and/or having a general practitioner come and visit patients (this applies only to private insurance holders, for it is included in the fixed fee general practitioners get for people who are insured at regional health care insurance organizations);
* costs related to the fixed fee for general practitioners;
* costs of prescription of drugs.

[86] The main difference between the model's depicted in Appendices 6 and 7 is that the model presented in Appendix 6 contains only the patients flows and influencing factors whereas the model depicted in Appendix 7 is the complete model as it was used during the participative policy modelling sessions. Note that some of the variables of the model have been given slightly different initial values in the model used to calculate the present chapter's outcomes. This to improve the quality of the graphical representation of the outcomes pf the runs discussed in the present chapter.

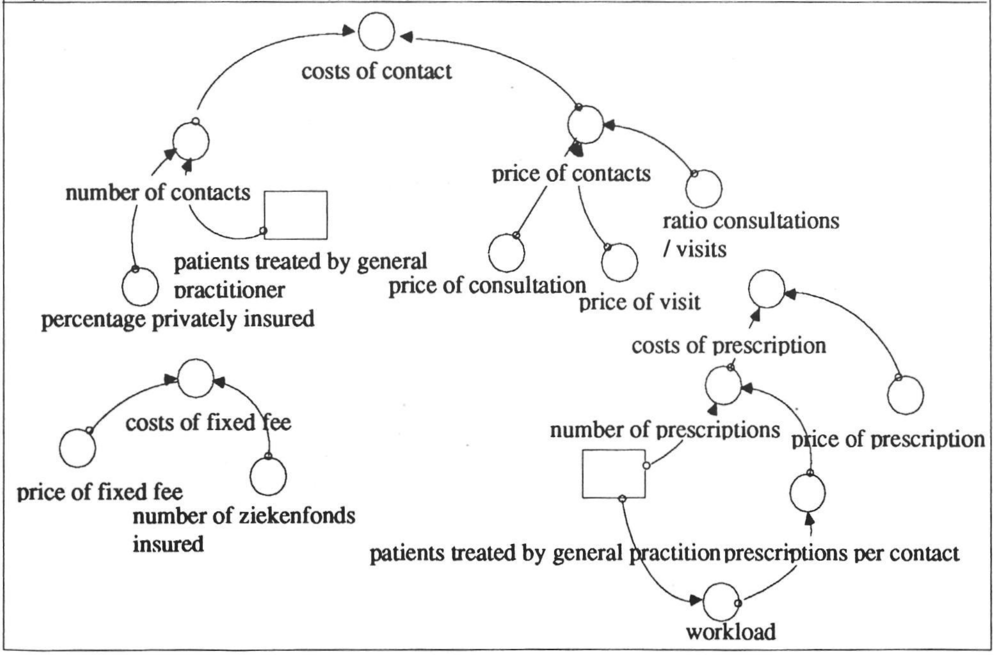
With respect to the costs of consulting a general practitioner or being visited by one, account must be taken of the percentage of people who hold a private insurance, the price of a consult, the price of a visit, and the ratio of consultations to visits, to arrive at a total representing the costs spent on having a contact (here contact means both a consultation and a visit) with a general practitioner. Note that a distinction is made between a consultation and a visit because visits (that is, having a general practitioner visit you at home rather than going to the general practitioner yourself) are more expensive than consultations are. Combining all three factors related to the price of a contact (i.e. both a visit and a consult), an average price of 25 Dutch guilders is arrived at for the year 1983.

The costs related to the fixed fees that are paid to a general practitioner for each person being insured at a 'ziekenfonds', can easily be calculated by multiplying the number of people who are insured at a 'ziekenfonds' in a particular year, by the amount of money that is paid per person. Since 100 Dutch guilder was paid per person in 1983, the costs spent on this aspect of the health care system can be represented as 100 times the number of people insured at a 'ziekenfonds'.

Regarding the cost of medicine, use can be made of the equation in which the average price of the drugs prescribed by a general practitioner is multiplied by the total number of drugs prescribed by general practitioners[87]. In 1983, the average price of a prescription (i.e. the medicine described on a prescription) was equal to 10.70 Dutch guilders.

The way in which the costs related to the general practitioner section are represented in the preliminary model can be depicted as follows:

Figure 5.14: Costs related to the general practitioner section




[87] Recall that the total number of drugs prescribed is a function of the number of contacts (consultations) a general practitioner has and the average number of prescriptions per contact (consultation), which, in turn, is affected by the general practitioner's workload. For an overview of the relationship between workload and average number of prescriptions, the reader is referred to Appendix 6.

### 5.5.3 MEDICAL SPECIALIST (OUTPATIENT VISITS)

#### *Patients flows*

In line with the description of the patients flows of the general practitioner section of the Dutch Health Care System, the presentation of the medical specialist section of the computer model, will be preceded by a description of the major processes that are included in the model. As far as the medical specialist section is concerned, patients are submitted to one of the following processes:

* Once the patients have been referred by the general practitioner to a medical specialist, they flow into the medical specialist section. The first thing they have to do, is to make an appointment to consult the medical specialist.

* On the whole, patients cannot be seen to by a medical specialist immediately. They first have to wait (on average about 4 to 5 weeks) before they can consult one. The reason for this is that in the Netherlands waiting-lists for a number of specialisms exist.

* Having spent some time in a level called 'patients waiting for the medical specialist', they do see the medical specialist, who can either carry out some medical transactions, or prescribe some drugs.

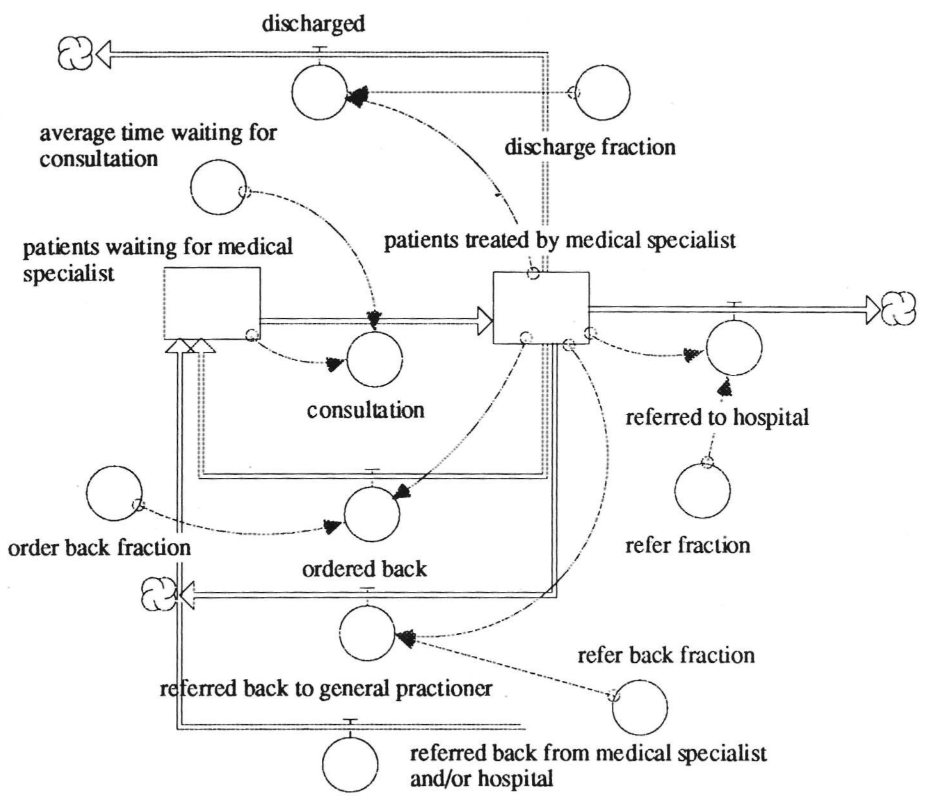
* Following the treatment, medical specialists have to decide whether the patient needs to be ordered back (for another treatment), referred to a hospital (for clinical treatment), referred back to the general practitioner (e.g. to have stitches removed), or discharged.

* Note that patients who are ordered back do not leave the health care system, but flow back to the level called 'patients waiting for the medical specialist'. To reduce the complexity of the model, no distinction was made between the time in between two consecutive consultations (i.e. the average order back time), and the waiting time for a first consult; the only waiting time that was built into the model, is the average waiting time related to the time that people are in the level 'patients waiting for the medical specialist'.

* Not all the patients who visit a medical specialist have come from either the general practitioner (referred) or the medical specialist him or herself (ordered back). Some of them have been referred back from hospital.

The patients flows corresponding to the processes just outlined, can be depicted as follows:

Figure 5.15: Patients' flows surrounding the medical specialist




To arrive at the initial values and parameter estimated included in figure 5.15, use was made of the annual reports of the hospitals that belong to the region of our client health care insurance organization, and the data bases of the client organization.

In the region of the client organization, 171 medical specialists are under contract. These medical specialists together treat about 8920 patients per week. Of these 8920 patients 17 percent is discharged, 2 percent is referred back to their general practitioner, 9 percent is sent to hospital (indicated for admission), and 72 percent is ordered back. The order back fraction of the medical specialists is substantially higher than the order back fraction of an average general practitioner: an average general practitioner orders back some 28 percent of the treated patients whereas a medical specialist has an order back fraction of more than 70 percent. This high order back fraction is one of the reasons why policy makers argue that the patients who flow into the 'expensive' right hand side of the Dutch health care system, have difficulty leaving the system. The reason why it is far more expensive to have patients who are ordered back in the medical specialist rather than the general practitioner section, is because medical specialists, in contrast to general practitioners, are paid per consultation (patients seen at the consulting-hours) and per medical transaction. To illustrate the difficulty of leaving the expensive medical specialist section, note that medical specialists have only a discharge fraction of 17 percent. Analogously to the level 'patients waiting for treatment by the general practitioner', the initial value of the level 'patients waiting for treatment by medical specialist' had to be inferred for no statistics were available on the number of people waiting for treatment. Multiplying the average waiting time (its value is based on a telephone-survey carried out by the client health care insurance company, and research carried out by the National Society of Medical Specialists (LSV, 1990)), by the order back rate flowing into that level, an initial value of 40.150 was estimated.

### *Influencing factors*

With respect to the factors that affect the patients flows surrounding the medical specialist, workload is expected to play the same vital role as it did for the general practitioner section[88]. Workload is defined as the number of patients treated per medical specialist per week. In line with the way in which general practitioners were expected to regulate their workload by changing their order back, refer and discharge fraction together with the order back time, medical specialists are expected to have the same tools at hand to regulate their workload. However, in addition to this, medical specialists can also vary the number of people they refer back from the hospital to the polyclinic (outpatients' clinic), for the same medical specialists often have both a consulting-hour, and work at the hospital. Since the way in which these feedback mechanisms operate is identical to the general practitioner section, they will not be described in great length in this section. For a more detailed description of the relationship between the medical specialist's workload and the variables it is having an impact on, the reader is referred to the overview given of the model in Appendices 6 and 7.

In addition to the impact workload is having on the patients flows surrounding the medical specialist, a relationship between workload and the average number of drugs prescribed per consultation and the average number of medical transactions carried out per consultation exists.

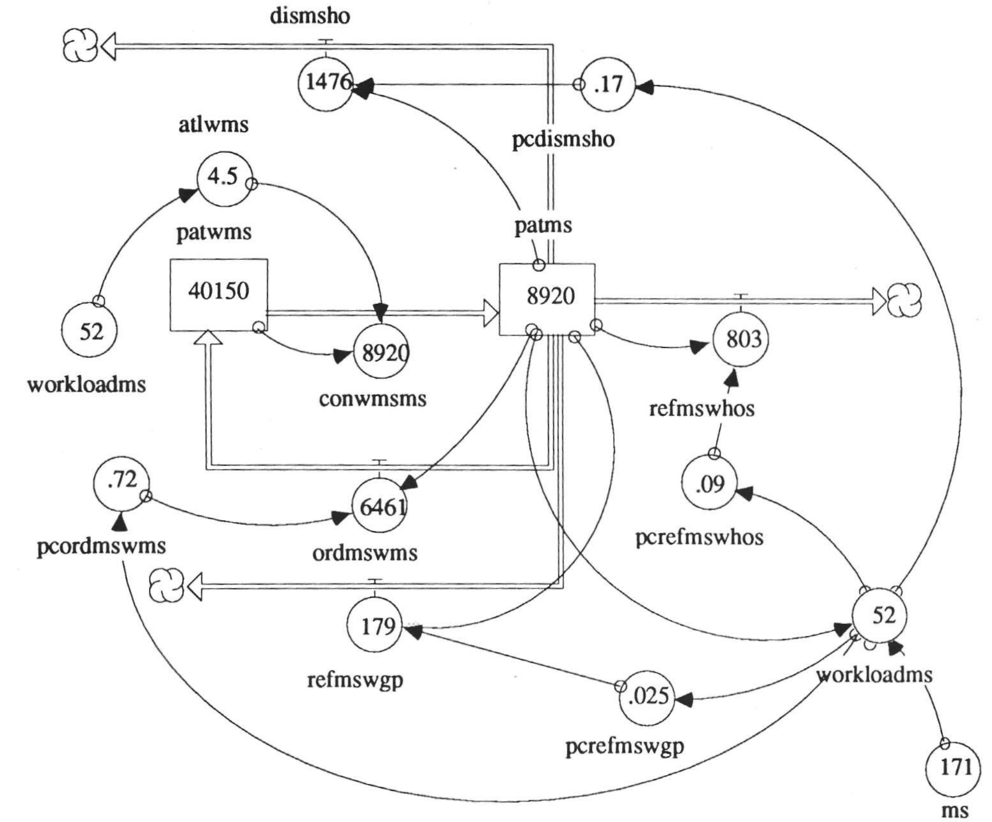
It is expected that the average number of drugs prescribed per consultation increases when medical specialists experience an increase in their workload, for the very same reason why general practitioners increase their average number of prescriptions when they face such an increase -prescriptions are used to end the consult, to shorten the amount of time spent per patient.

However, the relationship between the workload of a medical specialist and the number of medical transactions carried out by him or her is expected to be of an inverse nature: the higher the workload, the lower the number of medical transactions carried out during a consultation and vice versa. The reason for this is that the more patients a medical specialist has to see in the same amount of time, the less time (s)he has to carry out medical transactions. To be able to cope with this increase in workload, medical specialists confine themselves only to medical transactions that really need to be carried out. For an overview of the way in which the relationship between workload and average number of prescriptions and medical transactions is conceptualized in the model, the reader is referred to Appendices 6 and 7.

The impact 'workload' is having on the medical specialist section of the health care model, can be represented as follows (the numbers have been included to provide an idea of the quantities involved in this section of the model):

88 A correlation is found between the number of medical specialists and the volume of services offered by these medical specialists (KISG86, p. 57). Note that the number of medical specialists have been increased by 11.4 per cent in the period of 1983 to 1985. A plausible explanation for the correlation between the number of specialists and the volume of care is that additional medical specialists create their own work(load), as assumed in our model of the health care system.

Figure 5.16: Patients' flows and influencing factors




Levels:
patms = patients waiting for another treatment by the medical specialist
patwms = patients treated by medical specialist

Rates:
conwmsms = consultation (from level waiting for medical specialist to level medical specialist)
ordmswms = order back (from level treated by medical specialist to level waiting for medical specialist)
refmswgp = refer back (from level treated by medical specialist to level waiting for general practitioner)
refmswhos = refer (from level treated by medical specialist to level waiting for admission into hospital)
dismsho = dismissed (from level treated by medical specialist ot level people at home)

Auxiliaries:
ms = number of medical specialists
workloadms = workload of medical specialist
pcrefmswhos = percentage of patients referred from medical specialist to hospital
pcrefmswgp = percentage of patients referred back from medical specialist to general practitioner
pcdismsho = percentage of patients discharged
pcordmswms = percentage of patients ordered back for another treatment by a medical specialist
atlwms = average time waiting for a medical specialist

### *Costs*

The costs generated in the medical specialist section of the model are the following:

* costs related to consulting a medical specialist;
* costs related to medical transactions[89] carried out polyclinically by the medical specialist;
* costs related to the prescription of drugs.

The costs related to consulting a medical specialist have been calculated by means of the number and (average) price of the so-called referral notes. Recall that all the patients visiting a medical specialist for the first time need to have a referral note. Since the average price of a referral note can be calculated (67.70 Dutch guilders in 1983), the costs of first consultations can be determined. The people who visit a medical specialist for the second, third and.. time (these consultations are called repeated consultations [herhaalbezoeken], have to pay for their consultation as well[90]. The number of patients visiting without a referral note can be used as an indication for the number of these consultations, and together with the average price of a repeated consultation (27 Dutch guilders in 1983), the total costs of repeated consultations can be calculated. Taken together, they result in the costs spent on consulting a medical specialist.

With respect to the costs of medical transactions, multiplying the number of contacts (consults), with the average number of medical transactions per contact, and the average price of a medical transaction (82.50 Dutch guilders in 1983) will lead to a value of the money spent on medical transactions. Recall that the average number of medical transactions per contact is affected by the workload a medical specialist experiences.

Finally, the costs of prescriptions depend on the number of patients consulting a medical specialist, the average number of prescriptions per consult, and the average price per prescribed drug (in 1983, this was equal to 18.75 Dutch guilders).

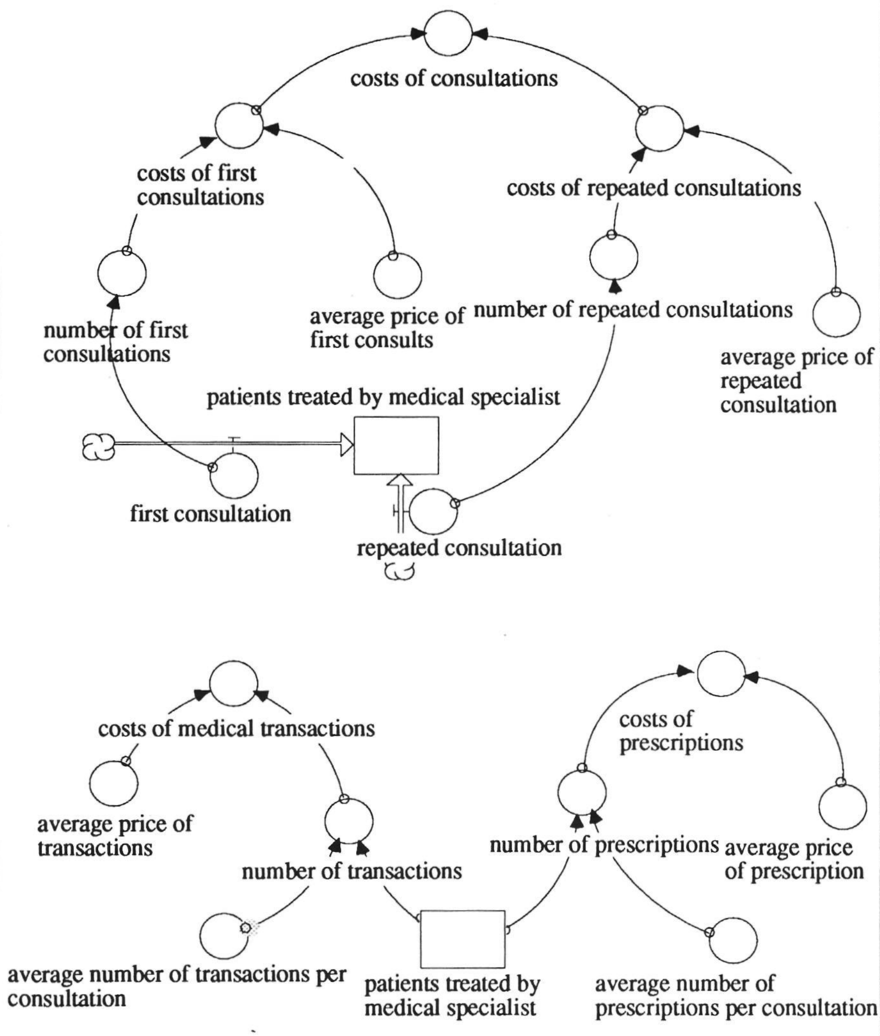
The above-mentioned three cost factors related to the medical specialist section have been represented in the preliminary computer model as follows:

---

89 The medical transactions included in the model concern the so-called polyclinical (outpatient) category III transactions [poliklinische tarief III verrichtingen] They are included because they contribute most to the costs of medical transactions.

90 In most cases, for if one is asked to return for a repeated consultation [herhaalconsult] within 1 month, the repeated consultation is not being paid for In 1990 a somewhat new system was introduced In the new system a distinction was made between so-called long and short referral notes, pointing to the period of time that they are valid, that is, can be used to visit a medical specialist

Figure 5.17: Costs of medical specialist section




## 5.5.4 HOSPITAL (INPATIENT ADMISSION)

Having outlined the basic elements of the general practitioner and medical specialist sections of the model of the Dutch Health Care System, account needs to be given of the third and final section: the section in which the hospital is predominant. In the description of the hospital section, patients flows will be focused upon first. This will be followed by an overview of the way in which the 'workload' of the hospital (the rate of occupancy) is having an impact on the patients flows surrounding the hospital, and the costs generated by this section of the health care system.

## *Patients flows*

Some of the patients who consult a medical specialist at a polyclinic or outpatients' clinic (the level 'patients treated by the medical specialist') are sent to hospital. Once they are sent to a hospital, they are submitted to the following:

* On the average, these patients have to wait some time before they can be admitted. There needs to be room on, for instance, a ward or an operation theatre, before they can actually be admitted, that is, flow into the level called 'hospital'.

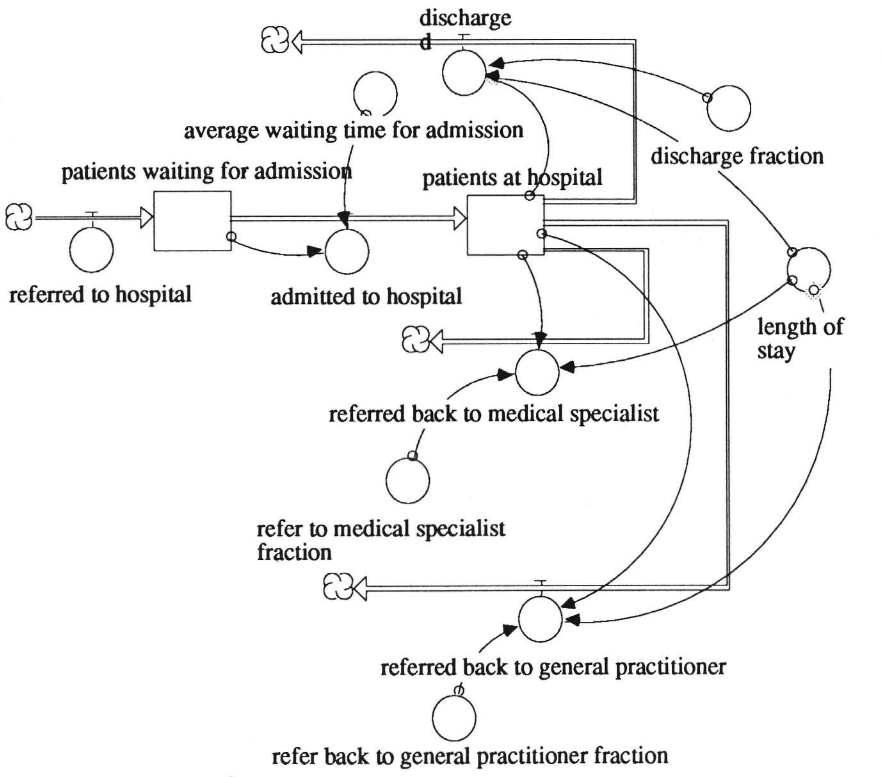
* Having entered the hospital, they spend some time on a ward, the so-called average length of stay.

* During their stay in hospital, some medical transactions are carried out or drugs are being prescribed.

* The patients who have recovered, are allowed to leave the hospital. They are either discharged (i.e. go home), referred back to a medical specialist (for polyclinical / outpatient treatment), or are referred back to a general practitioner.

The major flows of the third section thus are the following:

Figure 5.18: Patients flows surrounding the hospital section



To quantify the patients flows of this section, annual reports of the five hospitals of the region were used Based on these reports, it was found that on average about 1,300 people (that is, or 1,300 inpatient days [beddagen]) stay in hospital in a particular week, with an average length-of-stay of 11 4 days (1 623 weeks, for the model's time unit is a week) Moreover, the following values for the fractions belonging to the outflows were found a 'refer back to general practitioner fraction' of 2 5 per cent, a 'refer back to specialist fraction' of 90 per cent, and a 'discharge fraction' of 8 per cent To estimate the initial value of the number of patients waiting for admission into hospital (the 'level patients waiting for hospital'), use was made of the LSV (1990) research Based on their findings, it was decided to assign an initial value of 8 weeks to the average waiting-time for admission into hospital

### *Influencing factors*

The factor influencing the patients flows in the third section mostly, is the hospital's equivalent of workload, that is, its rate of occupancy Recall that workload was defined as the quotient of the number of patients dealt with and the number of providers Analogously to this, rate of occupancy is defined as the number of patients in hospital in a particular period divided by the number of patients that could be in hospital at a maximum In other words, rate of occupancy is the quotient of the number of inpatient days and the number of beds The rate of occupancy can be taken as a measure of the workload for the higher the rate of occupancy, the more patients have been admitted to hospital It is expected that hospitals (their management and medical specialists) aim at maintaining the rate of occupancy at a particular level[91], just like general practitioners and medical specialists do for reasons of efficiency and profitability Basically, to adjust the rate of occupancy, the number of hospital beds, or the number of inpatients can be changed Since medical specialist and management cannot easily decide to change the number of hospital beds, the number of available hospital beds has been taken as an exogenous variable Consequently, use must be made of the second option, focusing on the number of patients in hospital to change the hospital's rate of occupancy

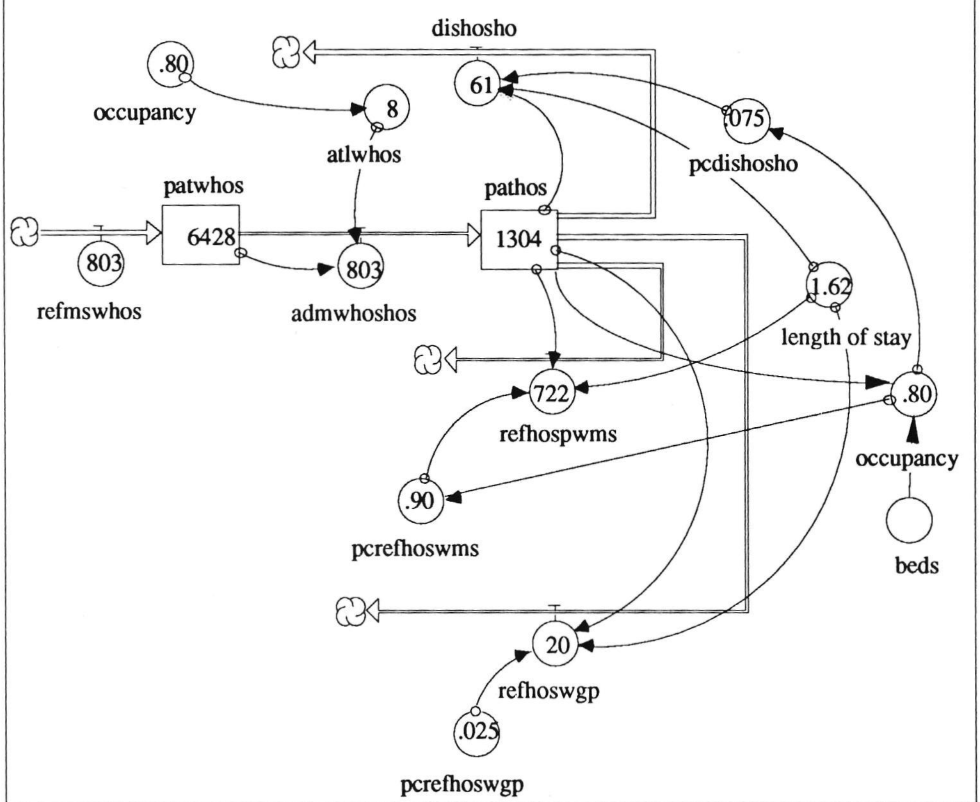
To change the number of patients in hospital, medical specialist can change both the average waiting time (the time that patients have to wait before they are admitted into hospital), and the refer back fraction by means of which patients are being sent from the level 'being treated by medical specialist' to a hospital[92] Lowering for instance, the refer back fraction, results in a reduction of the patients sent to hospital by the medical specialist because these patients leave the section in which specialistic treatment is provided (the patients for instance are sent to the general practitioner for control rather than being kept in the second or third echelon) Obviously, such a reduction in the number of patients sent to hospital is likely to lead to an increase in the number of patients ordered back for often these people need to be treated one way or another As said before, the exact nature of the relationships between the influencing factor and the patients flows is depicted in Appendices 6 and 7 For now it suffices to present the structural relationships between these two layers of the hospital section of the model of the Dutch Health Care System.

---

91 For the five hospitals of the region, an average rate of occupancy of about 80 per cent was found in 1983

92 In the first version of the model, average length of stay was also thought to be affected by the hospital s rate of occupancy However, since experts consulted on this issue stated that length of stay should be considered as an exogenous factor, affected primarily by developments concerning the technology used in hospitals, it was decided to remove the relationship between rate of occupancy and length of stay

Figure 5.19: Patients flows and influencing factors




Levels:
patwhos = patients waiting for admission to hospital
pathos = patients in hospital

Rates:
refmswhos = refer (from level treated by medical specialist to hospital)
admwhoshos = admitted (from level waiting for admission to level patients in hospital)
dishosho = discharge (from level patients in hospital to level people at home)
refhoswms = refer back (from level patients in hospital to level waiting for medial specialist)
refhoswgp = refer back (from level patients in hospital to level waiting for general practitioner)

Auxiliaries:
beds = number of hospital beds
occupancy = rate of occupancy
pcrefhoswms = percentage of patients referred back to medical specialist
pcrefhoswgp = percentage of patients referred back to general practitioner
pcdishosho = percentage of patients discharged
atlwhos = average time waiting in level waiting for admission to hospital
length of stay = average time spent in the level patients in hospital

## *Costs*

Regarding the costs generated at the hospital section, three sources have been discerned:

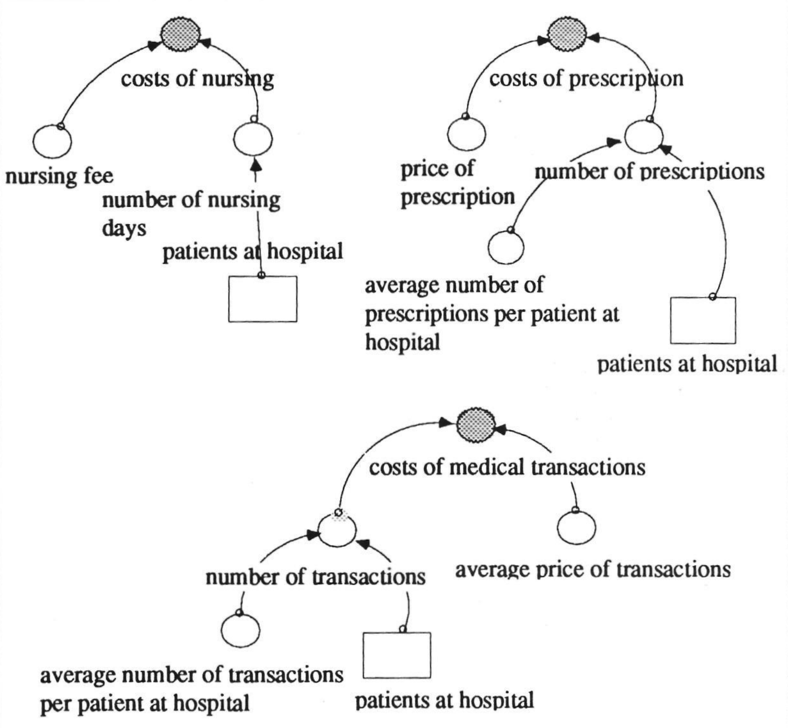
* costs related to nursing fees;
* costs related to medical transactions carried out clinically, that is, during the person's stay at a hospital;
* costs related to the prescription of drugs.

The costs of nursing are the product of the number of patients in hospital and the average price of a day in hospital. In 1983, the average price of a day in hospital was equal to 321 Dutch guilders.

The costs of medical transactions are based on the very same factors that play a role in the calculation of the costs of medical transactions in the medical specialist section. The number of patients in hospital, the number of medical transactions carried out per patient in hospital, and the average price of medical transactions (in 1983 this was equal to 177.50 Dutch guilders) thus lead to the costs spent on medical transactions.

Finally, the costs of drugs depend on the number of patients in hospital, the average number of drugs prescribed per patient in hospital, and the average price of drugs prescribed in hospital. Note that the average price of drugs prescribed was equal to 18.75 Dutch guilders in 1983. The factors that affect the costs generated by the hospital section can be depicted as follows:

Figure 5.20: Costs of the hospital section



## 5.6 AN OVERVIEW OF THE DUTCH HEALTH CARE MODEL IN ITS TOTALITY

### 5.6.1 INTRODUCTION

Having described the structure of each of the three sections that together make up the preliminary model of the Dutch Health Care System, account will be taken of the model's dynamic behaviour. One way to study the model's dynamic behaviour, is by disturbing the steady-state model's balance and see how it responds. To illustrate this procedure, and to give an indication of the dynamic behaviour of the model, an overview will be given of the effects brought about by a gradually increase of the number of patients consulting the general practitioner (the effects of a ten per cent increase in the consultation fraction will be displayed). Note that this particular test was used in the first session of the participative policy modelling program to illustrate the dynamic behaviour of the model. In addition to the description of the effects of an increase in a particular in-flow three kinds of runs will be described to give a more detailed impression of the dynamics of the Dutch Health Care System. In the first run, account will be given of one of the sensitivity analyses that were carried out to examine the robustness of the model - its sensitivity to changes. The second run will describe the effects of adding three exogenous variables to the steady-state model (the number of general practitioners, the number of medical specialists, and the size of the population). This description also serves to illustrate the nature of the activities carried out in the second session of the participative policy modelling method, for it was at the second session that participants were invited to consider the effects external influences would have on the behaviour of the model. Finally, in the third run, the dynamic responses of the model to a particular policy measure will be presented to provide some understanding of the dynamic properties of the preliminary computer model and to illustrate how the effects of policy measures can be examined by means of a system dynamics computer model.

### 5.6.2 DYNAMIC BEHAVIOUR OF THE HEALTH CARE MODEL

The model depicted in figure 5.21 consists of three sections: a general practitioner section, a medical specialist section, and a hospital section. Within each of the three sections, three layers have been discerned: a layer representing the system's patients flows, a layer in which the most important influencing factors have been included, and a layer representing the costs generated by that particular part of the health care system.

With respect to the dynamic behaviour of the preliminary model, note that wherever possible, historic values have been used to initialize the model's levels. However, in some cases estimations had to be made to arrive at an initial value for the level in absence of relevant statistics. Note that the values of the model were only slightly adjusted to arrive at the steady-state model represented in Appendix 6[93]. The initial values of the variables included in the model have already been presented in figures 5.13, 5.16, and 5.19. The abbreviations used in those three figures correspond to the abbreviations used in Appendix 6, in which the equations of the model are represented[94].

---

93 A steady-state model is a model in which the accumulations in a system are unchanging. A steady-state model is often used to examine the dynamic behaviour of the model because one only has to disturb the model's balance to see how it responds. Once people understand the behaviour of a steady-state model, they can proceed with examining the model's responses under more real-world conditions, for example by adding some important external influences. Another advantage of initializing the model in a steady-state is that it ensures that the system is in a state in which it still has a full range of response options open to changes to the model, rather than being constrained due to its initial state (Ithink User's Guide, 1991).

94 In figure 5.21, a graphical presentation of the model of the health care system is given. Note that the drugs prescribed section has not been included in that figure. This to avoid the figure from becoming to complex to read.

### 5.6.2.1 A GRADUAL INCREASE OF THE CONSULTATION FRACTION BY 10 PER CENT

Since the steady-state model itself does not show much variance in behaviour, all the graphs show horizontal lines (that's why it is called a steady-state model), an extra in-flow at the 'patients treated by general practitioner' level was given to the base-run to be able to examine how (well) the model behaves. To realize such an increase in in-flow, the consultation fraction was gradually, that is, in 5 year's time, increased by 10 per cent in total.

The effects of this change on the model's four major levels (people at home (patho), patients treated by the general practitioner (patgp), patients treated by the medical specialist (patms), and patients at hospital (pathos)) are the following:

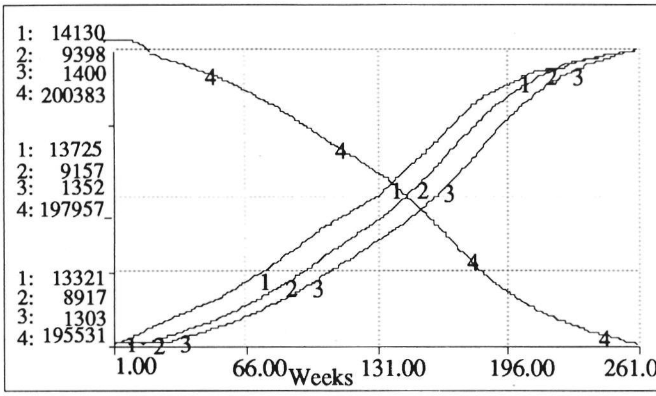



Figure 5.22:
Effects on four levels

As shown in figure 5.22, an increase in the in-flow automatically results in a decrease of the number of people that are not participating to the health care system, for the more people consult a general practitioner, the less people stay at home, leading to a decrease in the size of the level 'people at home'(#4).

Due to the increase in the number of patients treated by the general practitioner (#1), an increase in the number of patients treated by the medical specialist (#2) is brought about as well. The increase of the number of patients treated by the medical specialist follows the increase of the number of patients treated by the general practitioner, not only in size (although its increase is lower than the increase at the general practitioner, which makes sense for not all the extra patients treated by the general practitioner are referred to the medical specialist), but also in time in that the curve is more to the right. This is due to the waiting time in between the levels 'patients treated by general practitioner' and 'patients treated by medical specialist'.

The effects of this gradual increase of the consultation fraction on the general practitioner section of the health care system are depicted in figure 5.23.

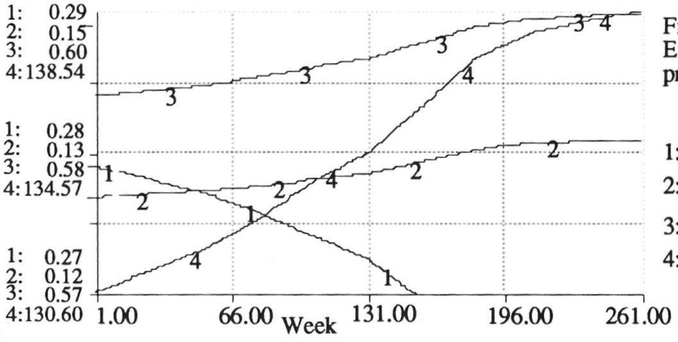



Figure 5.23: Effects on general practitioner

1: pcordgpwgp
2: pcrefgpwms
3: pcdisgpho
4: workloadgp

The figure shows that because of this increase in first consults, the general practitioner's workload (#4) is increasing. To compensate for this increase in workload, an increase in the fraction of patients referred to the medical specialist (#2) is brought about and the order back fraction (#1) is decreased. The discharge fraction (#3) seems to increase - relatively more patients are sent home.

Due to the increase in both the number of patients treated by the general practitioner and the refer fraction, the number of patients treated by the medical specialist increases as well. As depicted in figure 5.24, this will lead to an increase in the workload of a medical specialist (#1), as a result of which the fraction concerning the number of patients referred to hospital (#3) is increased and the medical specialist's order back fraction (#2) is decreased. Note that the average waiting time in the level 'patients waiting for the medical specialist' (#4) is increased because of this increase in workload. Again, it is important to examine the vertical axis, for it shows that both fractions change only very slightly. The curves, however, are useful in that they illustrate the direction of the change, even though the absolute value of this change is small.

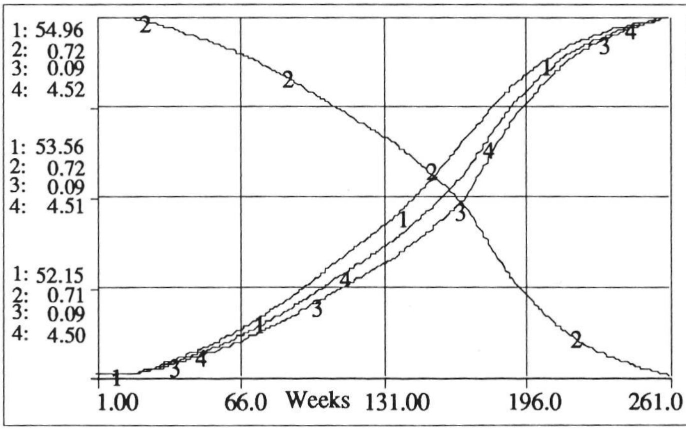



Figure 5.24: Effects on medical specialist

1: workloadms
2: pcordmswms
3: pcrefmswhos
4: atwms

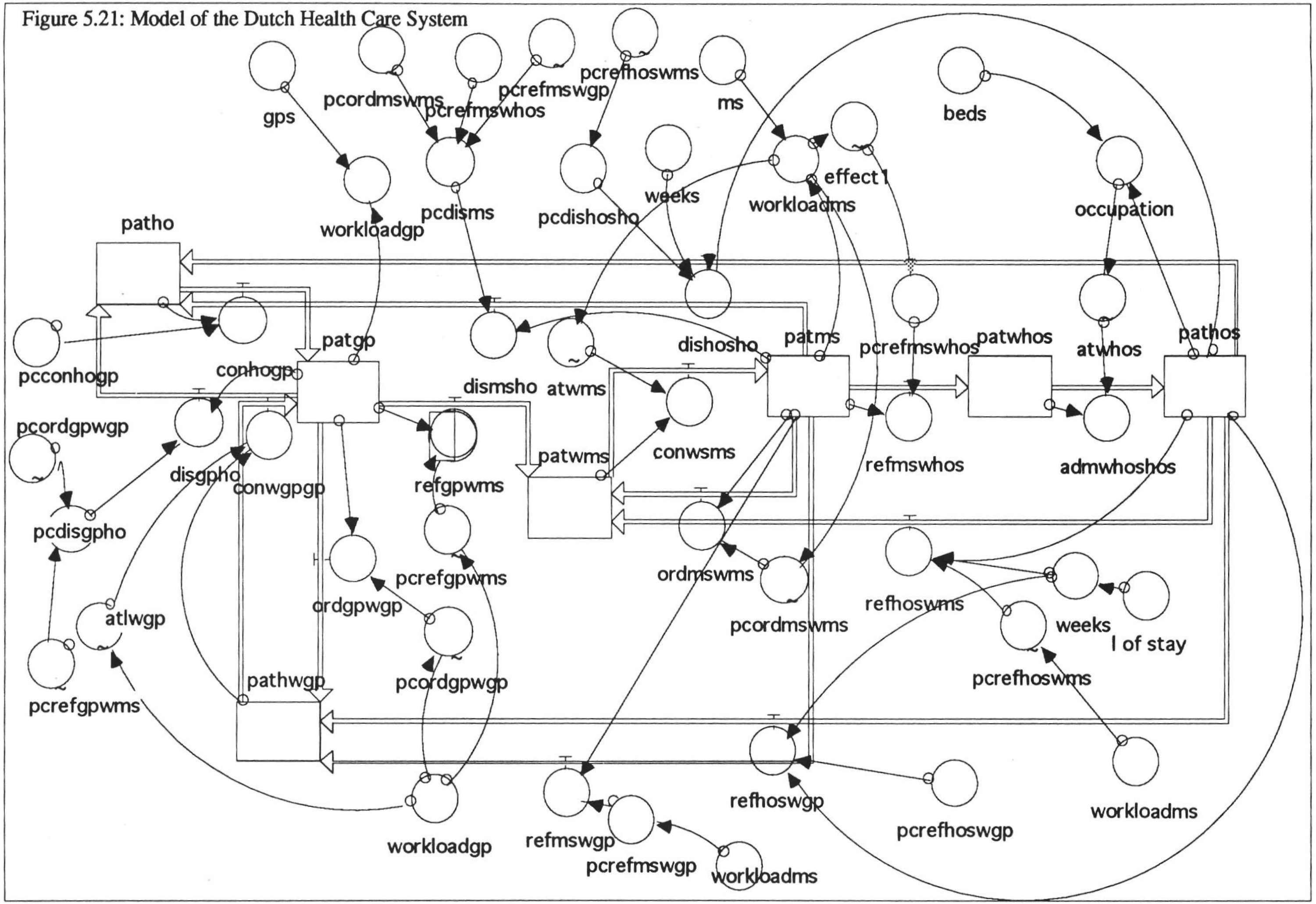


Figure 5.21: Model of the Dutch Health Care System

Model of the Dutch Health Care System

LEVELS

| | |
|---|---|
| patho | = patients at home |
| patgp | = patients treated by the general practitioner |
| patwgp | = patients waiting for another treatment by the medical practitioner |
| patwms | = patients waiting for treatment by the medical specialist |
| patms | = patients treated by the medicasl specialist |
| patwhos | = patients waiting for admissioninto hospital |
| pathos | = patients admitted into hospital |

FLOWS

| | |
|---|---|
| conhogp | = consulting a general practitioner |
| disgpho | = discharged by the general practitioner |
| ordgpwgp | = ordered back by the general practitioner |
| conwggp | = consulting the general practitioner for a second or ..th time |
| refgpwms | = referred by the general practitioner to the medical specialist |
| conwmsms | = consulting a medical specialist |
| ordmswms | = ordered back by the medical specialist |
| dismsho | = discharged by the medical specialist |
| refmswgp | = referred back from the medical specialist to the general practitioner |
| refmswhos | = referred to the hospital by the medical specialist |
| admwhoshos | = admitted into hospital |
| refhoswgp | = referred back from the hospital to the general practitioner |
| refhoswms | = referred back from the hospital to the medical specialist |

PERCENTAGES

| | |
|---|---|
| pcconhogp | = percentage of patients consulting a g.p. |
| pcordgpwgp | = percentage of patients ordered back by the g.p |
| pcdisgpho | = percentage of patients discharged by the g.p. |
| pcrefgpwms | = percentage of patients referred to the medical specialist |
| pcordmswms | = percentage of patients ordered back by the medical specialist |
| pcrefmswgp | = percentage of patients referred back from the medical specialist to the g.p. |
| pcrefmswhos | = percentage of patients referred to hospital |
| pcdisms | = percentage of patients discharged by the medical specialist |
| pcrefhoswms | = percentage of patients referred back from hospital to medical specialist |
| pcdishosho | = percentage of patients discharged from hospital |
| pcrefhospwgp | = percentage of patients referred from hospital to general practitioner |

MISCELLANEOUS

| | |
|---|---|
| workloadgp | =workload general practitioner |
| workloadms | = workload medical specialist |
| occupation | = occupancy rate |
| beds | = number of hospital beds |
| gps | = number of general practitioners |
| ms | = number of medical specialists |
| l of stay | = length of stay in hospital (in days) |
| weeks | = length of stay in hospital (in weeks) |
| atlwgp | = average time in level patients waiting for another treatment by the g.p. |
| atwms | = average time in level patients waiting for another treatment by the medical specialist |
| atwhos | = average time waiting before being admitted into hospital |

Finally, an overview needs to be given of some of the effects that are brought about at the hospital section of the model by this change in consultation fraction. In figure 5.25, it is shown that due to the increase in the number of patients treated by the medical specialist, more patients are referred to a hospital, as a result of which the number of patients admitted into hospital increases (cf. figure 5.22 as well), and the rate of occupancy (#1) is increased. To compensate for this change in occupancy, a change is brought about in the fraction of patients referred back to the medical specialist ((#3), to reduce the likelihood that those people eventually are referred to the hospital again. The increase in the number of patients admitted into hospital also leads to an increase in the number of patients waiting for admission, however, as shown in figure 5.25, no real increase in the average waiting time (#2) was brought about except for the last period of time where the average waiting time increased from 8 weeks to 8.02 weeks.

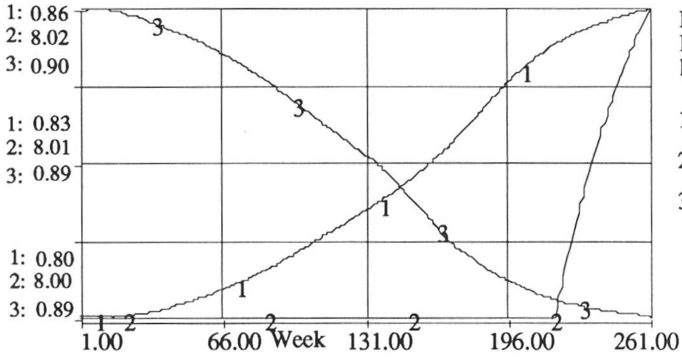



Figure 5.25: Effects on the hospital section

### 5.6.2.2 SENSITIVITY ANALYSIS

Sensitivity analyses are carried out to analyze the dynamic responses of a system to changes in the model. It serves to get an idea of how sensitive the model is to these changes and as to whether the model's sensitivity is in accordance with the sensitivity of the real-world's system. To test the robustness of the system, use is often made of logical functions (e.g. pulse, step, ramp etc.), initial values of levels are changed, and changes to the structure of the model are being made (e.g. adding or removing a feedbackloop) (Richardson & Pugh, 1981).

To illustrate how sensitivity analyses can be carried out, and to provide additional insight in the dynamic behaviour of the model, a brief description will be given of the effects of an idealized test-input: the effects of a pulse function on the rate which flows into the level 'people at home'. The pulse function that was used for this is PULSE (5000,0,75) representing an additional in-flow of 5000 persons in the level 'people at home', at times t=0, t=75, t=150, and t=225.

In figure 5.26, it is shown that indeed the number of people in the level 'people at home' is increased with 5000 every 75 DT (#4). Due to this increase, the number of patients treated by the general practitioner (#1), the medical specialist (#2), and the number of people admitted into a hospital (#3) increases as well. Note however, that because of the feedback mechanisms, the number of the patients treated by the medical specialist and the number of patients at a hospital increase much smoother than the number of people at home and patients treated by the general practitioner do.

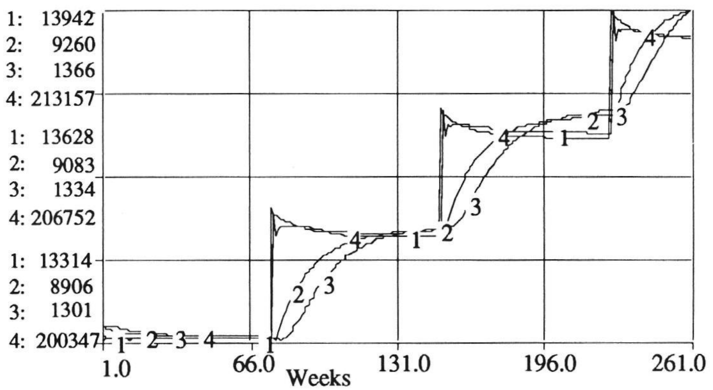



Figure 5.26:
Effects on four levels

1: patgp
2: patms
3: pathos
4: patho

With respect to the general practitioner section (see figure 5.27), it is important to note that due to the increase in the level 'people at home', the number of first consultations is increased as well, as a result of which general practitioners will experience a higher workload (#4). As shown in figure 5.27, general practitioners try to reduce their workload when confronted with such an increase in the number of first consults, although the actual reduction in workload is not impressive; only a small fraction of the pulse is compensated for.

The way in which general practitioners can compensate for such an increase in workload, is by reducing the order back fraction (#1), increasing the refer fraction (#2), and decreasing the order back fraction[95]. Note that the discharge fraction (#3) is increased as well: relatively more patients are being sent home.

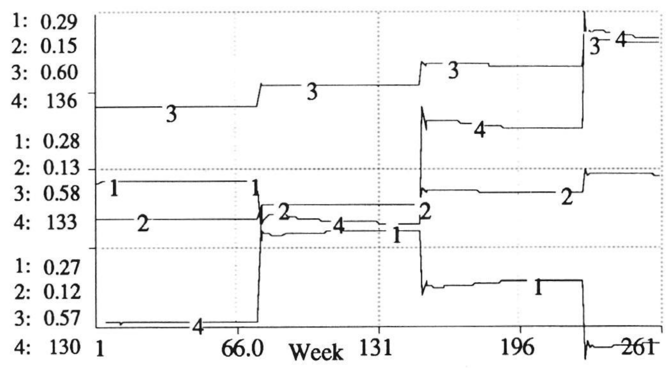



Figure 5.27:
Effects on the general practitioner

1: pcordgpwgp
2: pcrefgpwms
3: pcdisgpho
4: workloadgp

Regarding the medical specialist section, the curves have been smoothed, due to both the delays in between the sections of the model (waiting time), and the feedback mechanisms included in the model. In order to compensate for the increase in workload (#1), the fraction of patients ordered back is decreased (#2), and the fraction of patients sent to

95 General practitioners can also increase the time between consecutive consultations, that is, increasing the average time patients who are ordered back spend in the 'waiting' level. However, such a reaction only leads to a very temporarily result (which general practioners do realize) for those who are ordered back do come back for another visit. This 'measure' therefore is not very often taken by general practitioners to release some of their workload.

hospital is increased. The average waiting time in the level 'patients waiting for treatment by the medical specialist' (#4) is increased as well. It takes longer before patients can actually see the medical specialist. Note that in this section, each of the pulses is leading to a new equilibrium. However, just when this equilibrium is being arrived at, a new pulse is being given to the in-flow of the model.

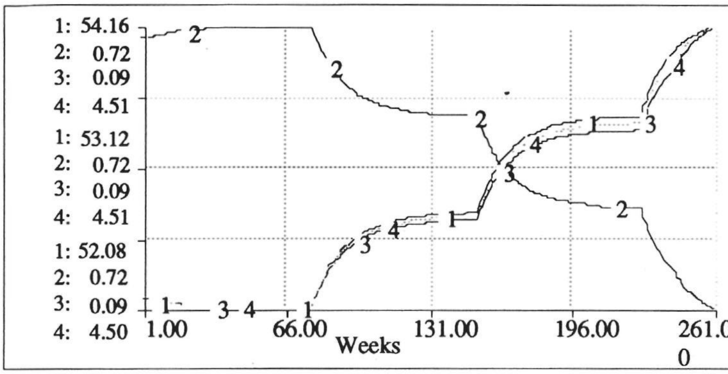



Figure 5.28: Effects on the medical specialist section

1: workloadms

2: pcordmswms

3: pcrefmswhos

4: atwms

Finally, regarding the section of the hospital, an increase in the average occupation rate is brought about. Note that the curves are even more smoothed in this section, due to the time it takes for patients to flow from the section in which the disturbance (the pulse function) was attached to the model to the hospital section, and the feedback mechanisms operating in the model. To stabilize the rate of occupancy (#1), the average waiting time is (#2) is increased slightly, and the percentage of patients referred back to the medical specialist section (#3) is reduced. The effects of the PULSE function on the hospital section of the preliminary model of the Dutch health care system can be visualized as follows:

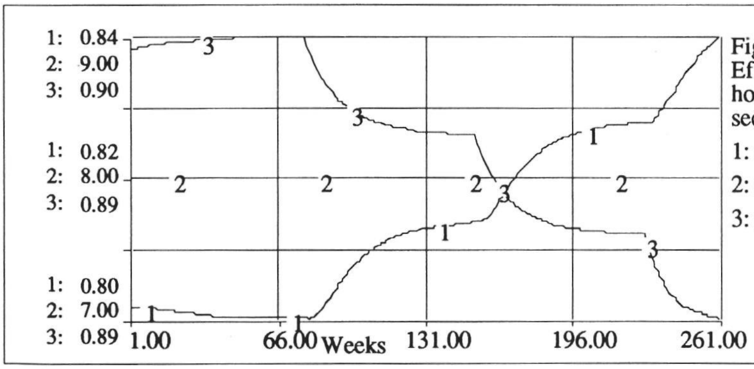



Figure 5.29: Effects on the hospital section

1: occupation

2: atwhos

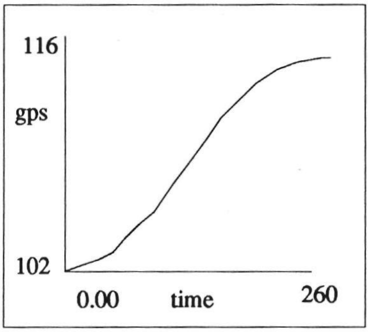
3: pcrefhoswms

### 5.6.2.3 EXOGENOUS VARIABLES

Next, to illustrate the dynamic behaviour of the model, a brief description will be given of the dynamic effects of three of the major exogenous variables. Another reason why the effects of the major external influences on the steady-state model are described in the present study, is to demonstrate what a more realistic health care model would look like, that is, a model that is not 'artificially' put into an equilibrium, or deliberately knocked out of balance as was the case with the sensitivity run described above. The exogenous variables that will be used for this concern the number of general practitioners, the number of medical specialists, and the growth of population. Their autonomous developments over time in the period 1983-1988 can be depicted as follows (figure 5.30 a,b,c):

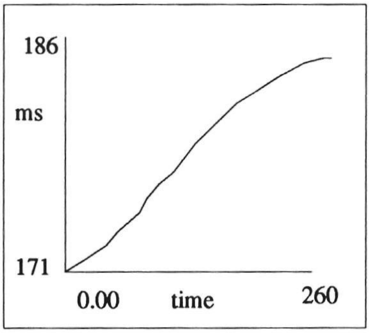
Figure 5.30: Exogenous variables

A: Exogenous variable: # of general practitioners



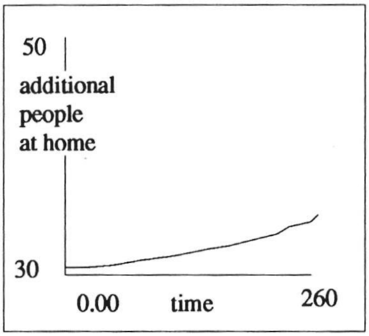

| input | output |
|---|---|
| 0.000 | 102.00 |
| 26.000 | 102.84 |
| 52.000 | 104.24 |
| 78.000 | 105.85 |
| 104.00 | 107.95 |
| 130.00 | 110.61 |
| 156.00 | 112.64 |
| 182.00 | 114.04 |
| 208.00 | 115.23 |

B: Exogenous variable: # of medical specialists




| input | output |
|---|---|
| 0.000 | 171.00 |
| 26.000 | 171.97 |
| 52.000 | 173.85 |
| 78.000 | 175.20 |
| 104.00 | 177.38 |
| 130.00 | 179.32 |
| 156.00 | 181.57 |
| 182.00 | 182.93 |
| 208.00 | 184.57 |

C: Exogenous variable: growth of population




| input | output |
|---|---|
| 0.000 | 30.35 |
| 26.000 | 30.25 |
| 52.000 | 30.70 |
| 78.000 | 30.80 |
| 104.00 | 31.10 |
| 130.00 | 31.40 |
| 156.00 | 31.80 |
| 182.00 | 32.60 |
| 208.00 | 33.30 |

The figures shows that a gradual but persistent increase in both the number of general practitioners and the number of medical specialists has taken place in the period between

1983 and 1988. This increase is not unlike the increase in the number of practitioners that has taken place in other regions of the Netherlands (KISG86, p.19, p.47). The number of general practitioners subscribed to the regional health care insurance company rose from 102 to 116, whereas the number of medical specialists under contract rose from 171 to 186. The graph representing the growth of the population in that particular area, is based on predictions made by the Dutch Central Bureau for Statistics [CBS][96]. The growth in population is about 30 persons a week.

To understand the combined effects of these three exogenous factors, one should bear in mind that basically two kinds of mechanisms are triggered by the changes. The first mechanism is activated by the growth of population, which is bringing about an increase in the number of people that use the health care system. The result of this is that an increase in the workload for both general practitioners and medical specialists is brought about. However, at the very same time, the number of practitioners is increased as well, and because this increase exceeds the increase in population, the patients to practitioner ratio is showing a decline. Due to the decrease in patient to practitioner ratio, a decrease in the workload of both general practitioners and medical specialists is brought about. To compensate for this 'lack of patients', the number of patients treated is increased somewhat - the treatment offered to the patients is intensified, patients are ordered back more often, and/or more medical transactions are carried out. The effects of this expansion, brought about by the mechanisms of growth in population and intensified treatment, on the number of people participating to the health care system, can be illustrated as follows:

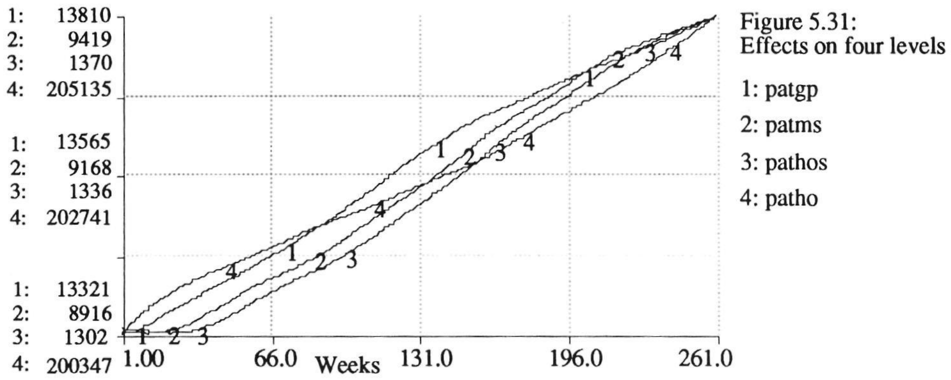



Figure 5.31:
Effects on four levels

The figure shows that the size of all four levels has increased. However, close examination of the figure reveals that the curves representing the levels 'patients treated by general practitioner' (#1), 'patients treated by medical specialist' (#2), and 'patients at hospital' (#3) are steeper, indicating that they grow more rapidly. The additional increase at those three levels can be attributed to feedback mechanisms that aim at stabilizing the workload in each of the three sections.

Figure 5.32 shows that indeed general practitioners order back more patients (#2) in an attempt to compensate for a decreasing workload (#4). Moreover, as explained in the description of the general practitioner section, they decrease the number of patients referred to the medical specialist (#3), despite the growth in population. It is the very same growth of population that is responsible for the rise in patients referred to the medical specialist starting at about t=140 weeks. At that time, the number of patients

96 As a matter of fact, the so-called 'middle trend' was used to project the growth of the population in that particular region (STG, Bulletin No. 20, June 1989, pp. 90-99).

treated by the general specialist has risen so much, that it compensates fully for the decrease in refer fraction, as a result of which the product of the two (the absolute number of patients referred to the medical specialist) starts to rise again.

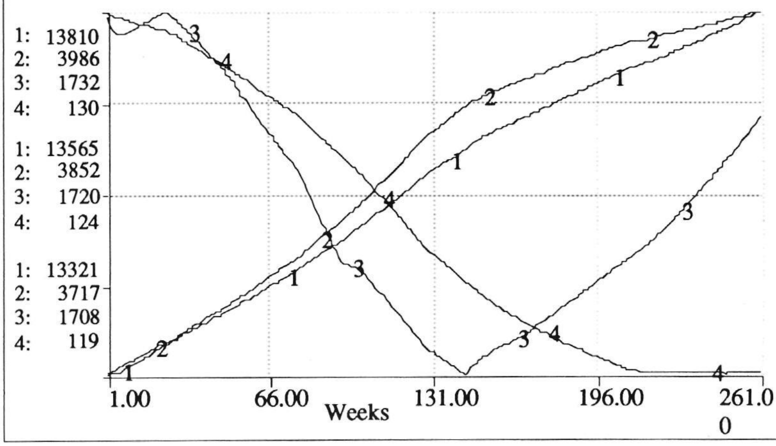



Figure 5.32: Effects on general practitioner
1: patgp
2: ordgpwgp
3: refgpwms
4: workloadgp

Regarding the medical specialist section, the increase in the number of patients treated by the medical specialist (cf. figure 5.31, curve #2) is the result of both the exogenous growth in population, and the medical specialist's response to a decrease in workload (figure 5.33, #1). A decrease in workload is leading to a reduction in the time people have to wait before they can visit a medical specialist (figure 5.33, #4). Moreover, to compensate for the decrease in workload, both the order back fraction (#2) and the fraction of people referred to a hospital (#3) are decreased as well, for the less patients are referred to hospital and the more are ordered back for another (polyclinical/outpatient) treatment, the more busy a medical specialist will be.

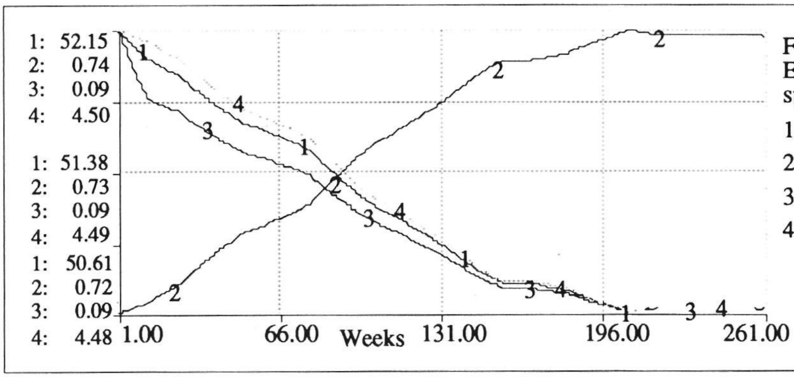



Figure 5.33: Effects on medical specialist
1: workloadms
2: pcordmswms
3: pcrefmswhos
4: atwms

Finally, account must be taken of the effects the exogenous variables have on the hospital section of the health care model (cf. figure 5.34). Not surprisingly, the rate of occupancy (#1) has increased substantially (recall, the population is growing). However, the response of this section of the model has been somewhat surprising for the fraction of patients referred back to the medical specialist (#3) shows an increase rather than the expected decrease. The reason for this is that the fraction is also affected by the workload of the medical specialist: a decrease in the workload is assumed to bring about an increase in the fraction of patients referred back to the medical specialist. Note that the average waiting time (#2) has not yet been affected by the increase in the number of patients in hospital. It is only when the rate of occupancy rises higher than its present value of 91 per cent, that an increase in the waiting time is expected; the rate of occupancy has not yet risen enough to bring about a change in the average waiting time. For an overview of the relationship between rate of occupancy and average waiting time, the reader is referred to Appendix 6.

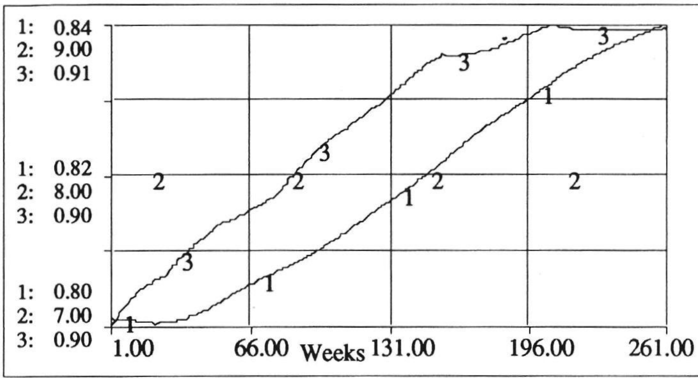



Figure 5.34: Effects on hospital section
1: occupation
2: atwhos
3: pcrefhoswms

### 5.6.2.4 POLICY MEASURES

Now that an idea has been given of the dynamic behaviour of the model of the Dutch Health Care System and the way in which the effects of sensitivity analyses and exogenous variables can be examined, account will be given of the way in which policy measures can be built into the model to study their dynamic consequences. Recall that the third session of the participative policy modelling session has been concerned with the effects of potential policy measures. Participants were given the opportunity to select a particular policy measure to examine their dynamic effects. The policy measure that will be focused upon in the present section concerns the number of patients admitted to hospital,[97] that is, referred by a medical specialist. It is expected that by having colleagues (medical specialists) from the very same specialism monitor occasionally each other's referral behaviour [intercollegiale toetsing], the refer fraction will decrease substantially (say, a reduction of about 10 per cent, that is, from .09 to begin with, to .08 once the measure has become effective).

It is expected that this measure will not be effective immediately after it is introduced, for time is needed to implement (that is, get used to) the measure. The effects of such a gradual decrease in the refer fraction on the four major levels of the model are the following[98]:

---

97 In the present chapter, it was decided not to include the costs and number of prescriptions and medical transaction because first of all they follow to a large extent the changes in the patient's flows, and secondly, because the major objective of the present chapter is to illustrate the dynamics of the model rather than giving a extremely detailed analysis of which policy measure is leading to which results in an attempt to arrive at recommendations on the basis of which the costs of health care can be reduced. Including factors concerning costs, drugs, and medical transactions would only complicate the matter unnecessarily. To acquire some understanding of the effects of policy measures on the ***costs*** of health care, the reader is referred to Gubbels, Verburgh, and Heine, ter (1992), and Verburgh, Gubbels, Vennix, and Post (1990).

98 Note that the policy measure is built into the model that was used to illustrate the effects of exogenous variables. In other words, to understand the effects of the policy measures, one should have to compare the effects presented in figures 5.35, 5,36, and 5.37 to the figures 5.30, 5.31, 5.32, and 5.33.

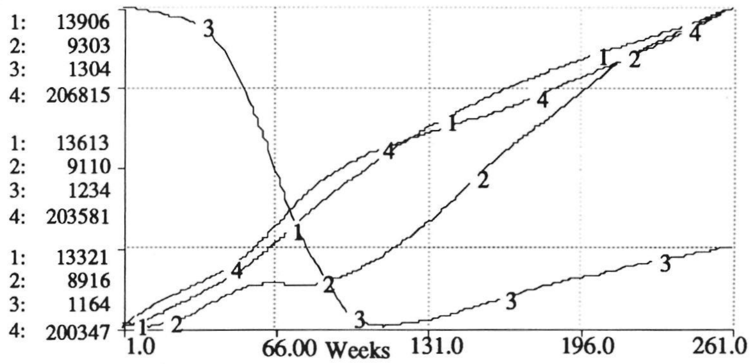



Figure 5.35: Effects on four levels
1: patgp
2: patms
3: pathos
4: patho

It is shown in figure 5.35 that indeed the number of patients in hospital (#3) has decreased, due to the reduction in the refer fraction. However, note that the curve is showing an increase in the number of patients in hospital after some time (despite the fact that the refer fraction is staying at the lower level), due to the growth in population. It is interesting to see that the number of patients treated by the medical specialist is showing a somewhat different behaviour, compared to the run in which the effects of three exogenous variables were depicted (cf. figure 5.31, graph #2). The rise in number of patients treated by the medical specialist (#2) of the present run levels off at some stage (and even stops for the time being, around time=66), in contrast to the run depicted in figure 5.31, where a constant rise in the number of patients treated by the medical specialist is found. Comparison of the two runs also shows that, due to the policy measure, more patients are sent home (#4), as a result of which the number of patients treated by the general practitioner is higher in the present run than it is in the run concerning the effects of the exogenous variables.

The major reason for this difference is, as is shown in figure 5.36, that the order back fraction (#2) cannot be higher than 74 per cent (maybe this should be changed in the next version of the model), so that the additional shortage of patients[99], due to the introduction of the policy measure, cannot be compensated for immediately. In other words, because the order back fraction cannot be raised unlimited, a fraction of the patients who would normally stay within the area where care is provided by medical specialists (either the level 'patients treated by the medical specialist' or the level 'patients at hospital'), will be leaving the section for they are no longer referred, and cannot all be ordered back. As to whether such an effect is realistic (or as to whether the assumptions of the model that bring about this dynamic behaviour still hold) can be made subject of discussion in the attempt to increase one's understanding of the problem.

99 The relative shortage in patients existed because the number of medical specialists has risen faster than the population has grown.

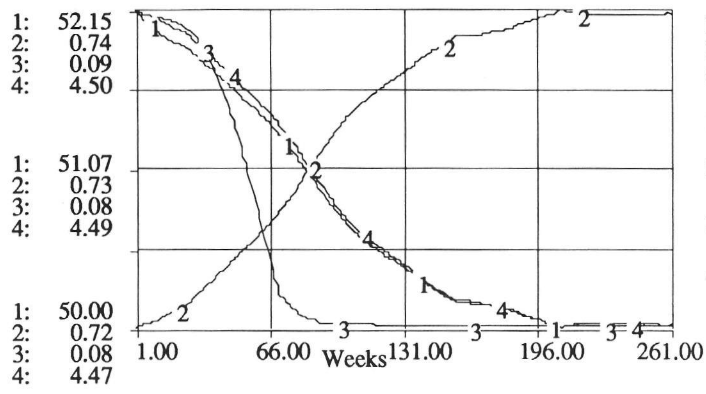



Figure 5.36: Effects on medical specialist

1: workloadms

2: pcordmswms

3: pcrefmswhos

4: atwms

Due to the reduction of the refer fraction (#3), the average waiting time for admission into hospital (#4) is reduced. The decrease in the number of patients in hospital (due to the reduction of the refer fraction, assuming that the policy measure will be effective) allows patients to be admitted into hospital faster (both the waiting time and the waiting lists will be reduced). Since it is assumed that only at a particular level of decrease in occupation rate (figure 5.37, #1) a change in waiting time (figure 5.37, #2) is brought about, and a kind of delay exists in the change of waiting time as well (for those who have already made an appointment, it is more difficult to change the waiting time), the curve representing the average waiting time follows the reduction in occupation rate not only in the shape of the curve, but also in time, as shown in figure 5.37.

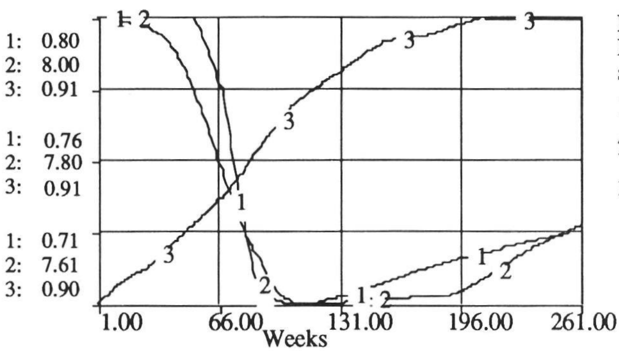



Figure 5.37: Effects on hospital section

1: occupation

2: atwhos

3: pcrefhoswms

## 5.7 VALIDITY AND STABILITY OF THE MODEL

Having discussed the formalization and quantification of the health care model used in the participative policy modelling sessions, and having given an overview of the dynamics of the system (describing the results of sensitivity analyses), account must be taken of another step in the model-building process not yet covered in the above description of the process and the product (model) resulting from it. The phase in the model-building process that needs to be dealt with now concerns the issue of validation, the degree to which the model is capable of representing the system under consideration, or the amount of trust that can be placed in the model. Following the discussion of the validity of the model, an answer will be given to the question raised in Chapter Two concerning the stability of the model to show that the likelihood of an error of the third kind can be reduced using the participative policy modelling method.

### *Validity*

With respect to the validity of the model, four classes of increasingly formal correspondence between the model and the information available about the simulant (the real-world system) have been discerned: common sense models, expert opinions models, partially estimated models, and fully estimated models (Randers, 1974).

In the common sense models (class 1 models), the model's assumptions are based on the modeller's intuition and general knowledge about the system. It is the kind of model that was created by Vennix et al. at the start of their project; a model created by the modeller's themselves on the basis of their assumptions of the process or problem they want to model.

In the 'expert opinion models' (class 2 models), by contrast, the model's assumptions represent the consensus of existing knowledge- as found in literature and among experts. These models also satisfy the requirements for class 1 models in being basically reasonable. To arrive of this class of models, experts are often invited out to participate in the model-building process. The model of the Dutch Health Care System presented in the present chapter seems to fit in this category of models, for not only was the conceptual basis created and agreed upon by quite a few experts in the Vennix study (Vennix et al, 1988), only few adjustments were made to the preliminary computer model by the experts taking part in the participative policy modelling sessions of the present study[100]. In addition to the intuitive assessment of the validity of a model (experts are asked whether they agreed or disagreed with the structure and/or behaviour of the model), the validity of the mnodel can also be assessed by means of empirical analyses. For this, use is made of the postdiction approach (in contrast to prediction) in which the correspondence between the model and the real world system is determined by comparing a simulated run (loaded with real-world begin-values) to some real world data. Comparison of the simulated behaviour of the number of patients in hospital in the period in between 1983 and 1988 with the real number of patients in hospital in that period, for instance, can be considered as an example of the use of the postdiction approach. The more the data generated by the model are in accordance with the actual facts (the real-world data), the more valid the model is. If no formal techniques are used (if only the *behaviour* is taken into account - the shapes of the curves of some of the model's variables are compared to the shapes of the curves of the behaviour of the real world system), the model resulting from these tests of validity will still qualify as a class 2 model. Note that the postdiction approach can also be used to create class 3 and class 4 models. However, rather than simply looking at the behaviour of the parameters (i.e. the shape of the curves they produce), numeric values need to be concentrated upon then.

Next in line in the number of classes of models that can be distinguished with respect to the degree of formal correspondence between the model's assumptions and reality, are the so-called partially estimated models (class 3 models). Partially estimnated models are models which have been validated on the basis of some formal techniques, that is, some formal techniques have been used to demonstrate the capability of the individual assumptions of the model to reproduce the real world data. Since no formal methods were used to assess the validity of the health care model, the model obviously cannot classify as a class 3 or class 4 model.

---

100 The only changes that were made, having been introduced to the preliminary model by means of the first workbook, concerned the introduction of a variable called 'operation theatre capacity' affecting the fraction of patients admitted into hospital, an increase in the value of the fraction of patients referred back from hospital to the medical specialist, and a change as a result of which the 'length-of-stay' became an exogenous rather than an endogenous variable.

The last category of models that can be created are the fully estimated models (class 4 models). In these models, the full system can be shown to be able to reproduce reality through formal techniques such as simultaneous equations estimation, and regression or Box-Jenkins' techniques.

Generally speaking, the confidence attributed to a model increases the higher the model's class number is. However, to move from a class 1 model (an intuitive common sense model) to a class 4 model (a fully estimated model), a lot of time and money is required. Moreover,because of the amount of time that is involved in the construction of a fully estimated model, chances are that the real world system under consideration has changed in the meantime. Both the problem's causal stratum and its political context may have changed at the time of the construction of the problem. The problems that may have to be faced in light of this, will be discussed in more detail under the heading of 'stability'. Another problem inherent to the construction of a fully estimated model is that more factors need to be included in the model to arrive at a greater correspondence between the model and the real world behaviour. This will lead to an increase in the complexity of the model, decreasing the model's transparency (Randers, 1974).

This is not to say that fully estimated models should not be built. However, it is meant to point to the fact that to determine the class of model to be constructed, account should be taken of the model's purpose: *"it is meaningless to judge validity in the absence of a clear view of the model purpose"* (Richardson & Pugh, 1981, p. 310). If the purpose of the model is to provide accurate predictions for particular variables (point prediction), than obviously the model needs to be a fully estimated one. However, if the model is constructed as a means to arrive at an understanding of the basic mechanisms underlying the behaviour of the system under consideration, then a completely different class of model would be appropriate. Hence it is recommended (Randers, 1974; Richardson & Pugh, 1981) to make use of the notion 'utility', rather than 'validity' to assess the value of a model. In this context utility is defined as the degree to which the model is capable of satisfying the *objectives* selected as important by the user, rather than the degree to which a formal correspondence between model and real world exists.

In the present study, the construction of the model was not carried out to arrive at a model which could serve as the Delphi oracle, predicting the exact consequences of measures to tell the organization what to do to decrease the costs of health care. By contrast, the preliminary model was constructed as a device that could be used by the participants to discuss their assumptions regarding the health care system in an attempt to arrive at an enriched and a shared understanding of the system under consideration. To start off such a discussion, the preliminary model does not have to pass the most rigid tests of validity (that is, it does not have to be a class 3 or 4 model). It suffices to use an expert opinion model, for not only do these models have a substantial amount of validity, they also ensure that the language used in the model, the concepts and relationships used to refer to the real world system, are part of the language spoken by the expert community participating in the process of adjusting and refining the preliminary model. It is important that the preliminary model exceeds the level of common sense models to avoid experts from feeling that their own level of expertise is not being called upon[101].

Note that as a matter of fact, the participants themselves did not want to be confronted with (and construct) a fully estimated model. The reason for this was that they were afraid that such a *"proven to be correct and thus true"* model would dictate the policy measures that had to be taken. It was only when it was made clear to the participants that the preliminary and revised models would not be regarded as the true or valid representation of the problem telling the organization what to do, but by contrast, would be used to

101 This is one of the reasons why we have found it extremely useful to have a liaison person at hand at all times during the project. The presence of such a liaison person allows one to check as to whether the model has the right amount of content expertise included. Moreover, it helps to acquire some understanding of the kind of people participating to the project, so that the program can be geared to the participants as much as possible.

improve both one's conceptualization of and communication about the problem at hand, that they were willing to participate[102].

So because the purpose of the model was to provide an increased understanding of the most important aspects of the complex health care system, and to discuss and communicate the individual participants' point of view in an attempt to arrive at a shared understanding of the problem, rather than using the model as a predictive device, the major validity checks that were carried out did not exceed the class 2 level in which validity is being assessed in an non-formal way. Consequently, the 'validation' questions that were asked and answered during the participative policy modelling sessions were questions such as 'Have the correct variables been included?', 'Is there a conceptual fit between the parameter values and the information available about the real system?', and 'Can the problem behaviour be replicated by the model?'. Moreover, because the purpose of system dynamics modelling, is to reduce the complexity of the real world system, complete correspondence between the real world system and the model should not be strived for, for this would inevitably lead to a model as complex as the real world system itself. In the context of the communication of knowledge and generation of insight, the right balance between simplicity and complexity, and aggregation and detail is probably even more important than the formal correspondence between the model and the real world system is.

Summarizing, in the context of the present study, a fully estimated model of the health care system seems to be inappropriate to strive for. Not only because the purpose of the project is to increase one's understanding of the problem, not to predict the future on the basis of the model, but also because system dynamics models represent the basic mechanisms underlying the problem only, to reduce the problem's complexity. In the process of conceptualizing the problem, behavioral validity and expert opinion validity seem to be sufficient and utility seems to be a more promising concept to use in the evaluation of the quality of the model, for the major purpose of the model-building process is to bring about an increased and shared understanding among the participants, rather than a careful prediction of the exact behaviour of the variables included in the model[103]. The evaluation carried out in the present study can be considered as an attempt to assess the *utility* of the model-building project carried out in the present study, that is, the degree to which it has succeeded in meeting its objectives, for as Greenberger (1976) puts it:

---

[102] Prior to the start of the project, two meetings had been organized to explain the purpose of the project. The first meeting was held with the organization's management team whereas the second was held to convince the two departments that they would benefit from participation to the participative policy modelling project. The objection regarding the status of both the preliminary and the revised model was raised in the second meeting.

[103] As said before, some attempts were made by the researchers to assess the model's consistency with reality using the postdiction approach However, it proved to be extremely difficult to acquire the information about the real world system required to compare the correspondence between the real world behaviour and the model of the system, hardly any statistics were available at that time on the health care production and consumption of that particular region despite that fact that all production and consumption is registered somehow by the client regional health care insurance organization It is interesting to note that the client is still in the process of making the potential valuable information they gather available for its management.

Moreover, some of the data that were available, concerned only one particular year (e.g Ree,van de, 1990), as a result of which they could not be used to compare the dynamic behaviour of the model with the behaviour of the real world system. The few comparisons that were made despite these difficulties have been reported in Verburgh & Gubbels, 1992 For the remaining variables, only conceptual assessment of the replicability of the real world behaviour was carried out, that is, experts were asked to state as to whether they agreed or disagreed to the behaviour of the model To illustrate, when confronted with the steady rise in costs generated by the model, experts participating in the construction of the model, were asked whether they agreed to the behaviour displayed by the curve Not surprisingly, all felt that indeed the costs of health care were represented accurately by means of the graph generated by the model

> "No model has ever been or ever will be thoroughly validated. ... "Useful", "illuminating", or "inspiring confidence" are more apt descriptors applying to models than "valid."

### *Stability*

The conclusion that the degree to which the model and the real world system correspond should be related to the purpose for which the model is built leads us to a refinement of the purpose for which the participative policy modelling project has been carried out. In the preceding section, it was stated that participative policy modelling aimed at bringing about a change in both the conceptualization of and communication about the problem at hand. In the present section, the notion of conceptual change will be expanded upon, by considering it from the point of view of the policy making process it is part of. This reconsideration of the change in conceptualization brought about by the policy making method brings us back to Chapter Two, in which the advantages of a system dynamics approach to the problem definition and conceptualization stages of the policy making process were discussed in great detail.

Recall that it was stated in that particular chapter that participative policy modelling could be used to prevent policy makers from making a serious errors when dealing with ill-structured policy problems: the error of solving the wrong problem. It was argued there that, to prevent people from making this so-called error of the third kind, more attention should be paid to the definition and conceptualization stages of the policy making process, and that people should be prevented from jumping to conclusions and/or defining the problem in terms of a solution they already have in mind.

To understand the relationship between making an error of the third kind and the notion of validity discussed in the preceding section, the difference between representing (and consequently solving) the correct problem and representing the problem correctly should be focused upon. The latter notion refers to the concept of 'validity', as discussed above, for it refers to the correspondence between the model and the real world system within one particular definition of the problem. The former notion, however, refers to the question as to whether one has the right problem in mind when starting the model-building process, and as such, is closely related to the so-called error of the third kind. To illustrate the difference in terms of the reference mode used in the model-building process: validity concerns the question as to whether there is correspondence between the reference mode and the behaviour generated by the model, whereas the issue of an error of the third kind is concerned with the question whether the correct reference mode has been used at the start of the project. As such it reflects the distinction between validity and utility, for reproducing the incorrect reference mode correctly will result in high validity but low utility for the purpose of the project - solving the problem - will not be met. Utility then, is used to refer to the degree to which the model has been successful in the process of solving a particular problem, that is, in the process of policy making.

With respect to the likelihood that people who are engaged in the participative policy modelling method commit a so-called error of the third kind, that is, solve the wrong problem and thus fail to meet the methods objective, it is expected that this likelihood is reduced substantially because people are compelled to spend much time on the conceptualization of the problem. As part of the participative policy modelling method, people first of all have to describe the problematic behaviour carefully (select and discuss a reference mode), forcing them to focus on the undesired behaviour stating explicitly what and why the behaviour of a particular variable is considered to be a problem (e.g. the 10 per cent rise in cost). Moreover, rather than jumping to conclusions, they first have to develop a causal network of variables that can account for the undesired behaviour, which allows them to, once more, define the problem very carefully. Finally, the consideration of the dynamic behaviour of the system forces those participating in it to reconsider their conceptualization of the problem in light of the feedback given by the computer model (see figure 2.10 for an overview of the way in which model-building stages are completed). This is not to say that participative policy modelling can avoid solving the problem at all times. However, it is mean to say that because of the many checks that are built into the method itself, the likelihood of a premature closure, leading to an incorrect definition of the problem and consequently to solving the wrong problem, can be reduced substantially.

Having said that the likelihood of an error of the third kind can be reduced if use is being made of a method that is taking the problem definition and problem conceptualization stages of the policy making model seriously, account should be taken of the fact that just *because* so much time is spent on these early stages, one still has to be concerned for the risk of solving the wrong problem. The reason why one, despite the use of participative policy modelling, has to be worried about the correspondence between the problem represented in the model and the problem that started the process of model-building, is because of the amount of time that is required to construct a model in relationship to the stability of the real world system - the world 'out there' is not waiting for us modellers to have the model completed but is changing at the very same time we aim at understanding it the way it is.

Two problems may occur when the real world system under consideration is not completely stable and is changing at the time of the participative policy modelling project. The first of these two problems is based on a narrow definition of the concept 'real world system', whereas the second one is based on a broad definition of what belongs to the system under consideration.

Defined narrowly, changes in the real world system concern only the causal stratum represented in the computer model. If the causal stratum of the problem under consideration has changed, and these changes have not been included in the model representing the system, the real world system's instability will lead to a decrease in both the model's validity and utility for not only does it lead to a decrease in the correspondence between the model and real world system (validity), it also affects the degree to which the outcomes can be used in the policy making process (utility)[104]. Note that the decrease in validity and utility brought about by the instability of the system under consideration is not related to the question as to whether the *right problem* has been solved (although it may well be that the problem no longer exists due to the changes in the real world system), for people may still feel that the very same problem should be dealt with. However, the solution selected on the basis of the model of the problem may no longer be valid, as a result of which the model may have become useless (utility) to arrive at solutions for the problem defined at an earlier stage.

To avoid this particular problem, several things can be done. Firstly, it is recommended that, wherever a choice exists, the most stable aspects of the system are focused upon. In the Vennix et al study (1988), this was done by concentrating on the patients flows rather than on the financial flows, for it was felt that irrespective of the changes in the financial structure of the Dutch Health Care System (recall that at that time it was expected that a dramatic change in the financial structure of the system was about to take place), the patients flows would remain the same for many years to come.

The decision to concentrate primarily on the patients flows has proven to be a good decision for the flows of patients within the health care system have not changed since the beginning of the project. The patients flows included in the model still hold today, about 4 years later; there has been no change in the way in which people move through the health care system. Moreover, representing the costs of health care as a derivative of the volume of the patients flows (together with the price of the product) still holds, despite the fact that some changes have been introduced regarding the financial structure, that is,

---

104 The relationship between validity and utility is such that an increase in validity is capable of bringing about an increase in the model's utility (not necessarily, as explained in the paragraph concerning the notion of validity), and vice versa. Utility by contrast, does not affect the degree of validity that can be attributed to the model. It may affect however, the degree of validity that is sought for.

regarding the way in which the money flows between the various actors of the Dutch Health Care System. Hence, the causal stratum of the health care system represented in the model has proven to be fairly stable[105].

Another solution to this problem, provided that the changes in the real world system are only limited, is to allow the participants to make some adjustments to the model. If the model is used for conceptual purposes, building these adjustments into the existing model need not be a serious problem, although unexpected behaviour may have a negative effect on the credibility of the model-building project[106]. However, when the model is intended to be used for predictive purposes, more serious testing and validation needs to be carried out before the adjusted model is able to meet its objectives, that is, serve as a 'valid' predictive device.

Finally, the risk of unanticipated (that is, not included in the model) changes in the real world system can be reduced by reducing the time spent on the model-building process. It is expected that, provided that two *experienced* model-builders work full time on the project, the project can be completed within 6 months rather than the 2 to 3 years it took in the present study. Note that an evaluation of the project is not included in this time estimate.

Regarding the problems that may occur if the stability of the system concerns a broadly defined real world system (as is the case with many social systems), note that in that case not only the causal stratum represented in the model may change, but the social-political context within which the problem has been defined and has acquired its shape as well. In the present project, changes that may have taken place in this decision-making stratum in addition to the causal stratum mentioned before (Klabbers, 1985) concern for instance the policy making context of the problem (e.g. 'Is the problem still an item on the political agenda or has it been removed from it?'). The instability of the system in this broad conception, thus may not only lead to an incorrect representation of the problem (an outdated model, including the wrong elements), but may also result in the representation of an incorrect problem, that is, a problem that no longer exists, or has been defined differently in the meantime. As such, participative policy modelling in general and policy making methods that take some time in general, run the risk of solving the wrong problem, that is, making an error of the third kind.

To avoid this particular pitfall, it is important to ensure that the participants do commit themselves to the definition of the problem on the basis of which the participative policy making process is carried out. If the client organization does not experience a particular problem, and subsequently agrees on a particular definition of that problem (e.g. a reference mode), chances are that they will change their mind at the time the model is being constructed. Moreover, having the policy makers take part in the process also helps to create the commitment that is required to have them still believe in the project and have it put on their political agenda, say, six months after the start of the project[107]. Obviously, the best way to deal with this particular kind of pitfall is to make sure that the project is carried out in a limited period of time. One of the objectives of the present study is to assess how the effects brought about by the participative policy modelling

---

[105] Note that in the period that the project was carried out (including the Vennix et al. (1988) study), several proposals had emerged to change the health care system. The plans proposed by the Dekker-committee (1987) were followed by the so-called 'plan Simons', after the Parliamentory Under-Secretary for Health Care [Staatssecretaris van WVC], which in turn was followed by the so-called 'Simons II' plan. However, only few of the changes proposed in these plans have been introduced yet. The health care system has proven to be a difficult to change system, due to, among others, differences in point of view among the many stakeholders.

[106] It is our experience that, rather than accepting the limitations of their own assumptions in considering dynamic effects generated by the model, participants tend to attribute unexpected behaviour to the model and the modelling method.

[107] Have client organizations pay for the project (even when it is only a small fee), will certainly affect their committment positively. Note that in the present study, no fee was being asked for.

method are affected by the way in which the method is carried out, and as to whether improvements can be made to the method (e.g. to speed up the process, increase the participants commitment etc.), for we are very well aware that due to the potential instability of social systems, errors of the third kind can never be completely ruled out if a lot of time is spent on the policy making process.

## 5.8 SUMMARY

Having given an overview of the organization taking part in the participative policy modelling sessions, and having presented the reasons why it was willing to take part in it, in light of the dramatic changes that were expected to take place in the Dutch Health Care System, an overview was given of the Dutch Health Care System itself and the preliminary model that was used to speed up the model-building process. Moreover, the description of the preliminary model served to illustrate what system dynamics models look like and what can be learned from analyzing the dynamic behaviour of such a model. Finally, the issues of validity and stability were focused upon, for they determine to a large extent the usefulness of the program offered to the participants.

In the next chapter, an overview will be given of the way in which the texts written by the participants have been transformed to arrive at the indicators needed for the construction of the present study's variables. With respect to the so-called cognitive mapping approach, the procedures followed to arrive at the indicators will be described in detail and the issues of validity and reliability that go together with these procedures will be addressed. Following this discussion, the way in which the indicators have been combined to construct the present study's variables will be presented and a summary will be given of the variables to be employed in the present study.

# CHAPTER 6: VARIABLE CONSTRUCTION

## 6.1 INTRODUCTION

In the present chapter, an overview will be given of the way in which the cognitive maps, on the basis of which the effects of the participative policy modelling method are being assessed, have been constructed. It will also discuss how the individual indicators have been combined to arrive at values for the variables that are needed to assess as to whether any changes in conceptualization have taken place.

With respect to the construction of the cognitive maps, an overview will be given of the procedures by means of which these maps have been arrived at, and the status of these cognitive maps in terms of validity and reliability.

Following the description of the construction of the cognitive maps, the construction of the variables will be dealt with, that is, an overview will be given of the variables that have been constructed on the basis of the questionnaire, and the pre- and posttest. In the description of the potentially specifying variables, account will be given to variables that concern the amount of time invested in the program, the background of the participants, and the participant-based evaluation of both the participative policy modelling program and the effects brought about by it. Following the description of the construction of the potentially specifying variables, a presentation will be given of the way in which the dependent variables have been constructed on the basis of the cognitive maps derived from the answers given to the pre- and posttests.

Finally, in the summary of the present chapter, an overview will be given of the variables that have been constructed and will be used to answer the research questions outlined in Chapter Three in the chapter following the present one.

## 6.2 CONSTRUCTION OF THE COGNITIVE MAP

### 6.2.1 CONTENT ANALYSIS

As explained in detail before, the present study aims to assess the changes brought about in the way in which people look upon a particular problem due to their participation in the participative policy modelling method. The changes in their conceptualization are examined on the basis of their verbal expressions, that is, on the basis of the texts written prior and posterior to the participative policy modelling sessions. To arrive at conclusions concerning some of the features of the structure and content of what is being asserted by the participants on the pre- and posttest, that is, with respect to the participants' discursive representations, the raw material (i.e. the written text) has to be reconstructed. The reason for this is that only on the basis of these reconstructed data (the cognitive maps) inferences can be made with respect to the features of the conceptualization or representation that are not directly observable but somewhat beyond the surface of the document constructed by the participants. In other words, the written texts have to be translated into cognitive maps to be able to draw conclusions that concern the conceptualization of the participants. These inferences are based on the theoretical framework that was used to determine what indicators were required to measure the theoretical constructs of strategic and domain-specific knowledge. This because, it is this theoretical framework that provides us with the clues that are needed to interpret changes in, say, the number of concepts, as a change in the domain-specific dimension of a person's conceptualization. The research process of making replicable and valid inferences from data to their context, that is, the process in which inferences are made within the context of a particular theoretical framework on the basis of a particular set of data, is called content analysis (Krippendorff, 1980). Although strictly speaking, the term content analysis can be applied whenever inferences are made on the basis of observable material to not directly observable entities, its use is usually restricted to the context of the analysis of documents and other forms of communication (Ridder, de, 1990). In the present study, content analyses are carried out because, in terms of Krippendorff (1980), observable material (the cognitive maps arrived at by a process called data making), is used to make inferences to arrive at conclusions with respect to a person's conceptualization within the context of a particular theoretical framework (the framework or analytical construct outlined in Chapter Three).

Since the raw material (the texts written by the participants on the pre- and posttest) has been converted (reconstructed) in order to arrive at the data on the basis of which inferences can be made regarding the major research questions of the present study, validity and reliability of the data making process will have to be examined.

With respect to the notion of reliability, the question is whether differences in the values found for each of the indirect indicators reflect differences in the written texts from which they have been derived, rather than resulting from the method (and person) used to transform the raw material into data to be used in the inference process.

The notion of validity first of all concerns the question whether the data selected to make inferences regarding a particular theoretical construct (e.g. the degree of domain-specific and strategic knowledge a person has), do indeed refer to the theoretical constructs to be covered in the present study. However, validity also plays an important role in the relationship between the written text and its corresponding cognitive map, for one certainly would like to be sure that the derived cognitive map is a valid representation of the discursive representation expressed by means of the texts written by the participants.

What these notions of validity and reliability exactly mean, and how they have been taken into account with respect to the content analysis method used in the present study to arrive at the features of a person's discourse on the basis of which an assessment is to be made of the degree to which people have changed their domain-specific and strategic thinking regarding the problem at hand, will be explained in more detail in one of the following sections. However, prior to discussing the content analysis method selected to assess the effects of the participative policy modelling method and elaborating upon the issues of validity and reliability, account should be taken of the features that need to be arrived at, *the reality of what the researcher wants to know about, the target for inferences* (Krippendorff, 1980, p. 172). It is only when we know exactly what the object of our research is, that we can decide what data and what analytical constructs are needed to be able to make inferences regarding that particular reality or target as Krippendorff calls it.

As to the present study, it has already been stated that its target (or purpose) is to examine whether people change their conceptualization of a particular policy problem, with respect to the *content* of their discourse (taken in an intuitive way, that is, content as representing facts and experiences (Krippendorff, 1980)), and its *structure* - the way in which the content is organized. The way in which these two kinds of changes in a person's conceptualization have been operationalized in the present study, has been explained in detail in Chapter Four, where an overview has been given of the indicators used to assess the degree to which people have changed their domain-specific and strategic knowledge concerning a particular problem at hand. Recall that this operationalization is based on the concepts used by the participants (both the kind and number of concepts that are used), and the (causal) relationships that exist among these concepts. The number of endogenous, exogenous and mono-disciplinary concepts for example, is used as an indicator by means of which the amount of domain-specific knowledge that is expressed in either the pre- or posttest can be assessed. To measure the amount of strategic knowledge included in the conceptualization of a participant, use is made of indicators such as the number and length of feedbackloops and the lengths of maximal paths.

Summarizing, the present study's target concerns the domain-specific and strategic dimensions of a person's conceptualization or discursive representation [redeneer-patronen]. The data that are needed to make inferences regarding the discursive representations or conceptualizations of participants (in the context of theories on policy making in general and coping with complex, ill-structured policy problems in particular), are data that concern the concepts and (causal) relationships included in the texts written by the participants on the pre- and posttest. However, to arrive at these concepts and relationships, the raw material will have to be transformed. In the next section, a more detailed description will be given of the cognitive mapping approach that is used to arrive at the indirect indicators needed to make inferences concerning the changes in conceptualization or discursive representation of the participants who deal with a particular policy problem.

### 6.2.2 COGNITIVE MAPPING

Having given an overview of the *data* required to make inferences regarding the effects of participative policy modelling on a person's discursive representation, account must be given to the (content analysis) *method(s)* that can be used to analyze the written texts so that the data described above can be arrived at, and inferences can be made regarding the target outlined above.

In order to arrive at the concepts and causal relationships that make up a person's discursive representation in the context of policy making, content analysis methods that focus on evaluative assertions (e.g. the evaluative assertion analysis (Osgood, Saporta, and Nunnally, 1956) or NET-Method (Cuilenburg, Kleinnijenhuis & De Ridder, 1988)) rather than the 'rational' causality hidden in the texts, seem less appropriate. The reason for this is that in the latter inferences are made with respect to the discursive representations expressed or incorporated in the texts rather than the writer's attitude towards particular objects. Consequently, it was decided to follow Vennix (1990) and employ the cognitive mapping approach developed by Axelrod (1976). This because the cognitive mapping approach is able to reconstruct text in such a way that the concepts and (causal) relationships included in it are made available for further analyses. For this the approach uses a construct called cognitive map. In addition to the close correspondence between variables needed for the present study and the variables constructed by means of the cognitive mapping procedure, it is claimed that, irrespective of the kind of research that is carried out, cognitive mapping is well suited to represent how policy makers look upon a particular problem: *"a cognitive map is a specific way of representing a person's assertions about some limited domain, such as a policy problem"* (Axelrod, 1976, p. 55). However, as stated above, the main reason why the cognitive mapping approach is adopted to reconstruct the raw material in such a way that inferences can be made regarding a person's conceptualization of a policy problem, is because of the close correspondence that exists between the data needed for this and the data arrived at by means of the method - concepts and (causal) relationships. Since *"...a cognitive map has only two basic types of elements: concepts and causal beliefs. The concepts are treated as variables, and the causal beliefs are treated as the relationships between the variables"* (Axelrod, 1976, p. 58), the use of the cognitive mapping approach to arrive at the variables needed to make the inferences we would like to make in the present study, stands to reason.

Now that has become clear that the cognitive mapping[108] approach will be selected to transform the raw material into cognitive maps on the basis of which inferences will be made regarding the changes in the participants' discursive representations (or conceptualizations) that take place due to participation in the participative policy modelling method, account will have to be given to the (data making) process by means of which the cognitive maps are constructed on the basis of the pre- and posttest measures.

---

[108]Cognitive maps that concentrate on concepts and causal beliefs are called 'cause maps' (Hall, 1989; Weick & Bougon, 1986)

### 6.2.3 THE CODING PROCEDURE

The coding procedure that was followed to arrive at a cognitive map follows Axelrod (1976) and Vennix (1990) closely. Basically, the procedure consists of reading the texts carefully to determine which of the statements of the text can be rephrased in terms of a causal relationship between the concepts used in the statements. Since only a few methodological guidelines were given by Axelrod to interpret different kinds of formulations, it was decided by Vennix to divide the coding process into several steps (Vennix, 1990; Vennix, Schmeets & Wester, 1987). Rather than coding the causal relationships directly, the coders in the Vennix study were asked to identify the concepts used in the texts and recode them on a separate codeform prior to establishing the relationships between these concepts. Moreover, in between the steps of identifying the concepts used in the texts and the establishment of the causal relationships between those concepts, the coders were asked to translate the concepts (wherever possible) into concepts from the external econometric simulation model. This to allow for an assessment of the degree to which the external model had been incorporated in the participants' conceptualization of the problem as a result of their participation. Once the relationships between the concepts had been established, the coders were asked to draw a cognitive map and to compare the cognitive map to the text it was based on, to check whether any changes to the map had to be made.

In the coding procedure that was followed in the present study, however, no distinction was made between the identification of individual concepts and the establishment of causal relationships. Since we knew we would be dealing with experienced coders, it was decided to speed up the coding process by identifying the concepts and establishing the causal relationships at the same time, using codeform A (cf. Appendix 8)[109]. While filling out codeform A[110], the coders were asked to stay as closely as possible to the vocabulary that was used in the text to avoid interpretation of the meaning of the concepts.

Once the causal relationships had been established, that is, once codeform A had been completed, the coders in the present study were asked to translate the concepts making up the causal relationships into so-called standard-concepts, wherever possible. The reason why the coders were asked to translate the concept from the text into a 'standard' concept rather than a 'model' concept, as was the case in the Vennix study, is to stress that an increase in correspondence to the external model is only one kind of improvement that can be made in the present study, and that a close correspondence with the external model is not the most important objective when dealing with ill-structured problems in the context of participative policy modelling. The reason for this is that it is more important in the context participative policy modelling that some kind of model-thinking is being arrived at, or that some kind of shared understanding is brought about, than that the external model is incorporated in one's conceptualization and that people have become alike in thinking in terms of the external model.

However, despite the fact that this study does not aim to assess the quality of the cognitive map in terms of correspondence with the external model, a translation process still has to be carried out in order to be able to compare the participants' conceptualizations *over time* (i.e. have they changed their individual conceptualizations) and *to each other* (i.e. have they acquired a shared understanding). If no translation is

[109]The distinction between scene and sentence employed in Codeform A, is used to distinguish between the question the answer is referring to (the number of scene), and the sentence within this particular answer (the number of the sentence). Since the are three questions, three scenes can be distinguished.

[110]For each of the relationships established, the coders were asked to attribute a value of direction. A positive value of direction (+) had to be attributed to the relationship in case the direction of the change was the same for both concepts (that is, the cause and the effect concept), whereas a negative value of direction (-) had to be attributed in case the direction of the change differed for each of the two concepts (e.g. an increase in the cause concept leading to a decrease in the effect concept or vice versa).

made, a change in terminology automatically is considered as a change in conceptualization, irrespective of whether the new concept (used on the posttest) has the same meaning as the old concept (used on the pretest). Using the concept 'computer' rather than 'PC', to refer to a machine with a keyboard, an internal memory device and a monitor, will be regarded as change in conceptualization if the texts written by the participants are examined on the basis of the exact terminology used in it only, rather than looking at the meaning of the concepts used, as well. It is important that the meaning (in an intuitive sense: denotation ánd connotation) underlying the concepts used in the pre- and posttest is changed before a change in conceptualization is to be concluded to. It is only when the concepts 'PC' and 'computer' are translated into one and the same concept (say, personal computer), that such an artificial change in conceptualization can be avoided.

Regarding the confounding effects that can be expected to take place with respect to the inter-individual changes in conceptualization in case no translation of the concepts used in the texts is carried out, note that artificial commonality may be arrived at as well. This because correspondence in use of meaning then is mistakenly restricted to correspondence in terminology only. Even when different words are being used, shared meaning may exist. An increase in correspondence in terminology thus not necessarily will lead to an increase in shared understanding, for the meaning of the concepts used will have to be taken into account as well.

To avoid artificial effects such as these, it was decided to construct a dictionary (Krippendorff, 1980, p. 125), in which concepts with the same meaning are put together in one and the same category, and have the coders translate the concepts used by the participants into standard concepts included in the dictionary as much as possible. The dictionary was created by the researcher on the basis of the concepts discerned by those who coded the raw material (cf. Appendix 11 for an overview of the dictionary that was created and used by the researcher). To carry out the translation process, the coders were asked to read the text once more, turn to the concepts of the causal relationships, and check which of them could be translated into one of the standard concepts included in the dictionary. The results of this translation process had to be written down on codeform B (cf. Appendix 9). Since the standard concepts were numbered, coders only had to write down the number of the corresponding standard concept in case a translation could be carried out successfully. In case the concepts could not be translated, however, the exact text had to be written down on codeform B.

Following the translation of the concepts, the sentences were read once more (this time by the main coder, that is, the person who coded all texts rather than just a sample of the texts written by the participants), to determine in which of the causal relationships a time-indication could be found. To qualify as a time-relationship, sentences such as 'will in the long term result in' or 'will immediately lead to', had to be added to the causal relationship between the concepts, because these qualifiers indicate that the author is aware of the fact that time is important in the examination of the effects brought about by a particular variable of the cognitive map. To code these time-indications, a (T) between brackets had to be written down in the comments box, that is, the box in which the coders were asked to place the sign of the relationship as well.

Moreover, despite the fact that quantifications do not appear on the list of indicators displayed in figure 4.11, the coders were asked to write down a (Q) between brackets in case reference was being made to a quantity. For instance, if it was stated that a decrease in the average length-of-stay (the measure proposed in question three), would be leading to a five per cent increase in the number of patients admitted into hospital, the concept 'number of patients admitted into hospital' had to be qualified as a Q(uantified)-concept[111].

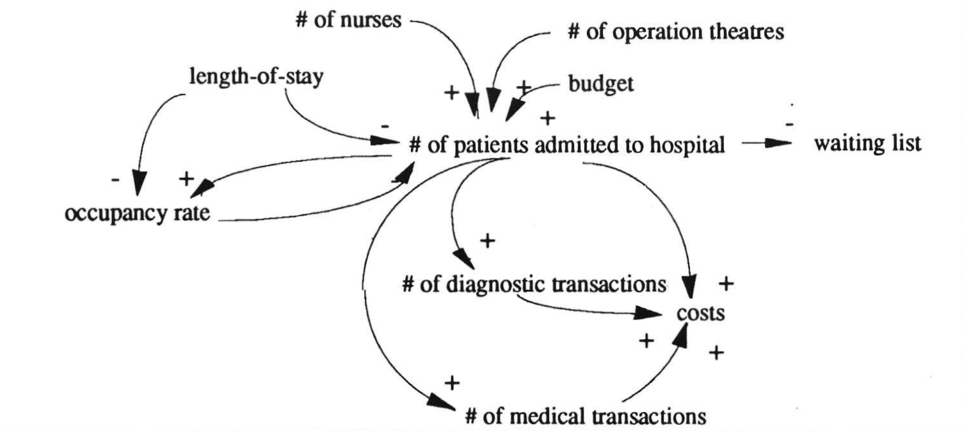
[111]In order to be able to determine the number of exogenous and endogenous concepts, all concepts were classified as either a 'general practitioner' concept, a 'medical specialist' concept, and a 'hospital' concept. Since this classification was carried out by the researcher independently of the coding process, and was

To summarize, the major steps of the coding procedure followed in the present study to transform the written texts into cognitive maps are the following (for a more detailed description of the guidelines, the reader is referred to Appendix 10):

* Read the entire text (that is, read the entire scene related to one particular question of the pre- and posttest).
* Read the text sentence by sentence and determine for each of the sentences whether one or more causal relationships can be detected.
* Attribute a sign (value) to the relationship (+ or -).
* Translate the concepts used in the texts into standard concepts.
* Determine which of the relationships qualify as T(ime) relationships, and which of the concepts qualify as Q(uantified) concepts.

In order to illustrate the coding process, look at the way in which a fragment of a policy note written to answer question number three of the posttest, has been recoded to arrive at a cognitive map. The codeforms (A and B) that have been used for this are presented in appendices 8 and 9. Note that they have only been filled out partly, just to illustrate how the procedure works. The text and corresponding cognitive map are the following:

Figure 6.1: Sample cognitive map




Answer to question III

*"Reduction of the length-of-stay will, in the short run, lead to an increase in the number of patients admitted into hospital (probably not more than 5 per cent), as a result of which the waiting list can be reduced. However, this obviously also depends on the number of operation theatres available, the number of nurses available, and the available budget.*

*Another effect of the measure is that the occupation rate will decrease, as a result of which the number of patients admitted to hospital will, after some time, increase (those with a less serious disease will be admitted to hospital as well to make up for the decrease in rate of occupancy). Because of this increase in the number of patients in hospital, the number of medical transactions and diagnostic transactions carried out will be increased as well, as a result of which an additional increase in the costs of health care will be brought about."*

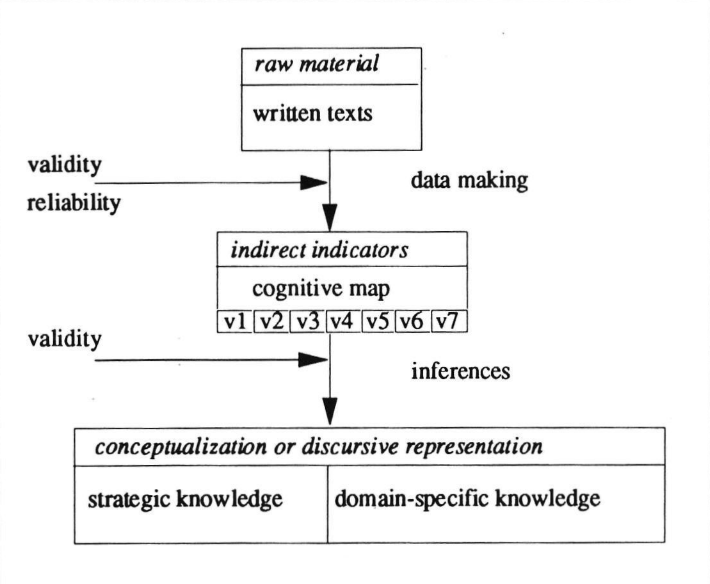
applied to the data automatically, that is by means of computer (the SPSSX program), it is not considered as a separate step in the coding process. For an overview of the list of concepts used by the participants to answer the questions asked on the pre- and posttest, and the way in which they have been classified to arrive at a kind of dictionary, the reader is referred to Appendix 11.

In figure 6.1 it is shown that only one feedbackloop is found in the text written by this particular participant: the direct feedbackloop between the number of patients admitted into hospital and the occupancy rate. Also clearly visible is that in the text written by our sample participant, use is made of exogenous variables such as the number of nurses, the number of operation theatres, and the budget (note that a true expert should have included a relationship between the costs generated in the hospital and the available budget as well), to account for changes in the number of patients admitted into hospital.

Having described the *reasons why* the texts written by the participants had to be recoded, the *method* that was selected for this (the cognitive mapping approach), and the *procedures* that were followed to arrive at these cognitive maps, account must be given of the validity and reliability of the cognitive mapping method. To do so, a description will be given first of the validity issues that can be discerned with respect to the data making (data construction) procedure and the inferences made on the basis of the cognitive map. This will be followed by a discussion of the reliability of the data making procedure and the steps that have been taken in the study to arrive at a value for the various forms of reliability that can be discerned.

In figure 6.2, an overview is given of the process by means of which the gap between the research questions and the raw material has been bridged. It also serves to make a transition to the issues of validity and reliability because it shows where the issues of validity and reliability fit in, in the overall process of transforming the raw material into material on the basis of which inferences can be made regarding a particular context (the context of the effects of participative policy modelling on the individual and inter-individual changes in conceptualization of a policy problem).

Figure 6.2: From raw material to research variables



### 6.2.4 VALIDITY

With respect to the validity of the present study, two kinds of validity will be discerned on the basis of the content-analysis process depicted in figure 6.2.

The first kind of validity concerns the validity of the data making procedure that is used to transform the written texts into cognitive maps. The major issue with respect to this kind of validity is the question whether the cognitive map can be considered as a valid representation of the texts written by the participants and whether the text is a valid representation of what is in the participants' mental models.

Regarding the second kind of validity, the validity of the inferences made on the basis of the indirect variables to the context of the present study, that is, to the strategic and domain-specific aspects of the participants conceptualization, the main question is whether the variables constructed on the basis of the indirect variables do indeed measure what they are supposed to measure: domain-specific and strategic knowledge.

To answer the first question, that is the question whether the cognitive mapping approach is a valid way of representing the concepts and relationships 'hidden' in the texts written by the participants, reference will be made to researchers who claim to have applied the cognitive mapping approach successfully to arrive at a condensed but valid representation of the way in which policy makers perceive a particular problem. Axelrod (1976, p. 7) for instance, claims that *"the validity of the documentary coding method is not yet fully established, although the empirical studies that use it offer different kinds of evidence in support of its validity"*. Moreover, applications of the cognitive mapping approach reported by Carley (1986), Eden (1988), Georgantzas & Madu (1990), Montazemi & Contath (1986), and Vennix (1990) illustrate that cognitive maps are currently used to represent the way in which people look upon particular issues. It thus seems that, as long as care is taken of the procedures by means of which cognitive maps can be derived, cognitive maps can be considered as valid representations of the way in which people look upon a particular policy problem. To ensure that the content of the cognitive map can be considered as a valid representation of the text written to answer the pre- or posttest, the coders were given an exercise book, in which the rules by means of which the raw material could be transformed into cognitive maps, were described in some detail. This to ensure that the correct procedures were followed in an appropriate manner, to arrive at cognitive maps that can be considered as accurate representations of the texts on which they are based.

In addition to the careful description of the procedures to ensure that the cognitive maps arrived at do represent the texts written by the participants accurately, account has been taken of the reliability of the procedure, for it is acknowledged that low reliability scores do affect the validity of the data making procedures negatively. In the next section, a more detailed description will be given of the issues that are related to reliability. However, prior to that, the validity of the inferences made on the basis of the cognitive maps to the context of the present study, will be elaborated upon.

As said before, the second kind of validity is concerned with the question whether the right kind of variables have been included in the present study. In other words, do the variables constructed on the basis of the indicators measure the theoretical constructs they intend to measure? To answer the question regarding the validity of the inferences made on the basis of the cognitive map variables, several approaches can be taken.

In the first place, some evidence of validity can be arrived at if the theoretical framework outlined in Chapter Three on the basis of which the variables have been selected and constructed is found to be sound in theory, because many of the variables follow directly from it. The question whether recursive paths do indeed measure feedbackloops, for instance, is becoming an academic one, once it is accepted that feedbackloops are an important element of strategic knowledge.

Secondly, some evidence of validity can be acquired by relating the operationalization described in Chapter Three to some other studies where more or less identical inferences have been made with respect to the content and structure (format) of how people look upon a problem. In this context, references should be made to the studies carried out by Vennix (1990) and Axelrod (1976), for a remarkable resemblance is found between many of the variables used to assess the (change in) content and format of the way in which policy makers look upon a particular problem[112] in these studies and the variables used to assess the content and format of one's conceptualization in the present study.

Having discussed the issue of validity in some detail, attention will be paid to the reliability of the coding procedure to assess the degree to which the instrument used to transform the raw material into data that can be used to make inferences to the context of the present study, will bring about the same results (that is, the very same cognitive maps), if carried out by different coders, at different points in time (Krippendorff, 1980). The relationship between reliability and validity is such that *"reliability is a necessary but not a sufficient condition for validity"* (Krippendorff, 1980, p. 129). Reliability does set limits to the potential validity of the research results, in that an unreliable research process can never result in valid conclusions. However, reliability does not guarantee the validity of the research results because *"two judges with the same prejudices may agree on what they see but be wrong by all other standards"* (Krippendorff, 1980, p. 129). Consequently, it is important that both the validity and reliability of the research process carried out in the present study are taken into consideration. Since the issue of validity has already been dealt with, account will be given of the issue of reliability in the next section.

### 6.2.5 RELIABILITY

#### *Types of reliability*

According to Krippendorff (1980), at least three different types of reliability can be distinguished: stability, reproducibility, and accuracy. They all concern the replicability of the procedure, that is, the degree to which replication of the procedure is resulting in the same outcomes.

Stability is defined as the degree to which a process is invariant or unchanging over time. It becomes manifest under test-retest conditions, that is, when the same coder is asked to code a set of data twice. Disagreement between the two ways the units are coded, reflects the so-called intra-observer inconsistency or noise. It is considered the weakest form of reliability for it only measures the degree to which one and the same person can replicate the outcomes of the process.

If, by contrast, one would like to be sure that the procedure can also be carried out by other persons in such a way that the very same results are being arrived at, reproducibility should be used, for reproducibility indicates the degree to which a process can be recreated under varying circumstances, at different locations, using different coders. To establish reproducibility, use is made of test-test conditions, that is, two or more individuals are asked to apply the very same recoding instructions on the same set of data. Disagreement among the coders reflects intra-observer inconsistency and inter-observer differences in the way in which the instructions are interpreted.

Finally, the strongest reliability test is called accuracy. It measures *"the degree to which a process functionally conforms to a known standard"* (Krippendorff, 1980, p. 131). To arrive at a score for this, test-standard conditions are being used, for instance, by comparing the performance of one coder to what is known to be the correct performance measure. Differences between the coder and the standard, reflect intra-observer inconsistencies, inter-observer disagreements, and systematic deviations from a standard. As such, it is closely related to validity if the standard is considered as the true or valid outcome of the process. However, since no such standard was available in the present

[112] In the next chapter, it will be shown that with respect to nine variables correspondence exists between the variables used by Vennix to measure the content and format of a conceptualization, and the variables used in our study.

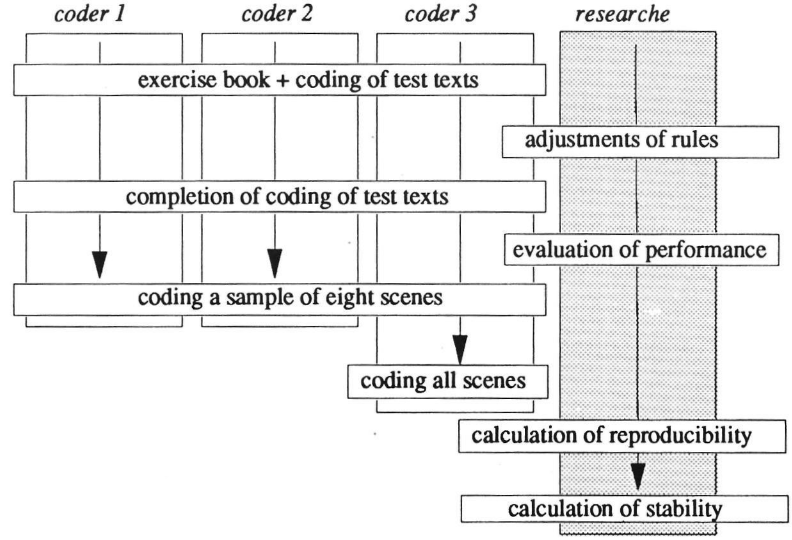
study, it was decided to follow the advice given by Krippendorff (1980)[113], to assess the stability and reproducibility of the process if accuracy scores cannot be arrived at.

The steps that were taken to assess both the stability and reproducibility of the coding process carried out to construct a cognitive map will be described next. An overview will be given of the way in which various coders have been involved in the coding process to arrive at scores on stability and reproducibility, the reliability measures that have been used to assess the degree of reliability, and the indicators of the cognitive maps for which reliability measures have been calculated.

### *Procedure to arrive at reliability scores*

To arrive at a score on stability and reproducibility, use has been made of the following coding design:

Figure 6.3: Overview of the coding process




As depicted in figure 6.3, most of the coding work has been carried out by the researcher himself. However, to determine the stability and reproducibility of the coding procedure, two other persons (i.e. coder 1 and coder 2) were asked to code a sample of the written texts as well. Note that coder 3 and the researcher are one and the same person.

To start with, all three coders were given an exercise book in which not only the rules that had to be used in the coding process were outlined, but also examples were given of potentially difficult sentences, to increase the likelihood of a correct application of these rules. Together with the exercise book, all three coders were given a sample text on the basis of which they had to fill out codeform A. In other words, at this stage, the coders were only asked to read the text sentence by sentence to determine whether one or more causal relationships could be detected in it. In addition to that, they were asked to attribute a sign to each of these relationships. It is important to note that at this stage the

---

113 *"But in most situations..., the standards against which the accuracy would be established are rarely available... Data should at least be reproducible, by different researchers, at different locations, and at different times, using the same instructions for coding the same set of data"* (Krippendorff, 1980, p. 132).

coders were not asked to start translating the concepts, for the dictionary that is needed for this, had not yet been constructed (because its construction is based on the completion of codeforms A). The completed codeforms A were not only used to construct the dictionary needed to complete the coding procedure, but also served to assess how successfully the rules were applied by the coders, and to determine whether any adjustments to these rules had to be made. Careful examination of the codeforms A revealed that, although no changes to the rules themselves had to be made, some additional information had to be given to ensure that all rules would be applied in a correct manner. Once the additional information had been added to the instructions, the coders were asked to once more check their completed codeforms A, to check whether any changes to their coding outcomes had to be made in light of the improved instructions, and to subsequently complete the coding procedure by translating the concepts into standardconcepts. For this, they were asked to use codeform B.

Evaluation of the codeforms B showed that sufficient agreement in the interpretation and application of the rules among the coders could be arrived at. Consequently, all three coders were given a sample of eight scenes (i.e. answers to a particular question) to be coded, using both codeforms A and B. The coders were given the very same scenes in order to be in the position to calculate the inter-observer (or inter-coder) reliability, that is, to arrive at a reproducibility score (Krippendorff, 1980).

Following the process of coding this sample of eight scenes, *all* scenes (162 in total: 27 participants completed both pre- and posttests) were coded by the researcher in a random order. Note that the researcher (coder 3) had also been coding the above mentioned sample of eight scenes, so that a score on his stability could be calculated as well.

Having outlined the way in which scores on stability and reproducibility have been arrived at, two more issues need to be dealt with in relation to the reliability of the procedure by means of which cognitive maps have been constructed.

The first issue concerns the elements for which a reliability score has to be constructed. Usually, in content analysis coders record single elements, such as the presence or absence of a phenomenon. To calculate the reliability, one simply focuses on these single elements. In the present study, by contrast, many elements had to be recoded (e.g. concepts, relationships, signs). Following Vennix (1990), it was decided to include only those elements in the reliability analysis, which act as an indicator in the theoretical model outlined in Chapter Three. As a consequence, reliability scores have been calculated for the concepts and the relationships incorporated in the cognitive map. Note that no reliability scores are calculated for the coding of the quantifications, time-indications, field of expertise, and endogeneity or exogeneity. The reason for this is that the cognitive operations that have to be carried out for this are minimal as a result of which they are carried out efficiently and reliably (Krippendorff, 1980, p. 63).

The second issue concerns the kind of measure that has been used to express the stability and intercoder reliability (reproducibility). Most measures take into account the fact that accidental agreement, on the basis of change, has to be controlled for in the eventual reliability score. In particular when the number of categories used by the participants in their classification endeavors is limited, expected agreement is something to control for, for it may result in an artificially high reliability score (for a more detailed description of the effects of accidental agreement due to a small number of prespecified categories on a reliability score and how to control for it, the reader is referred to for instance, Knottnerus & Volovics (1989), Krippendorff (1980), and Popping (1983).

In the present study, however, the number of categories that can be used to recode the written texts to arrive at cognitive maps, is almost infinitely. Not only does the dictionary used to translate the concepts used in the texts contain 252 categories, the actual number of categories that can be selected is much higher, because a coder can decide to create a new category if (s)he feels that the concept used in the text cannot be translated into a standard category. As a consequence, the chance that two coders accidentally (that is, by chance) code the same concept in one and the same category can be neglected. With respect to the relationships to be coded, the number of possible categories is even higher, for the 252 concepts included in the dictionary, result in 252 X 252 categories[114], that is potential relationships. Hence it was decided not to use reliability measures such as Cohen's kappa (Popping, 1983), and Scott's pi (Krippendorff, 1980), but to take the observed agreement P(o) as the measure for stability and reproducibility, ignoring the effects of any expected agreement P(e). In line with Vennix (1990), observed agreement has been defined as the *number of agreements between two coders divided by the total number of observations*. However, we differ from Vennix in that the total number of observations is not considered to be equal to the highest number of observations of both coders, but is regarded as being equal to the total number of differing observations made by the two coders. To illustrate this difference, consider the outcomes of a coding process, in which concepts have been coded by two coders (A and B):

Figure 6.4: Measuring reliability: an example

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A | c1 | c2 | c5 | c9 | c12 | c17 | c22 | c34 | c52 |
| B | c1 | c2 | c9 | c22 | c66 | c77 | | | |

coder A: 9 concepts
coder B: 6 concepts

Using the Vennix definition to calculate the reliability score with respect to the concepts coded in the example given in Figure 6.4, a score of .44 is obtained, for the number of concepts both coders have in common is equal to 4, and the total number of concepts, according to Vennix, is equal to 9. However, in this study, a reliability score of .36 would be arrived at, because the total number of concepts is considered to be equal to 11 rather than 9. The example thus clearly shows that the reliability measure used in the present study can be regarded as a more rigorous test of reliability than the Vennix measure. The reason why it was decided to use the more rigorous measure, is because the Vennix measure is not able to distinguish between the situation in which coder B codes 'c1, c2, c9, c22, c66, c77' (depicted in figure 6.4), and another situation in which the concepts 'c1, c2, c9, c22' are coded. According to the Vennix (1990) measure, both series would lead to the conclusion that the two coders (A and B) have 4 concepts in common, on a total of 9 concepts. In the present study's measure however, the number of differing concepts is also taken into account, as a result of which a different value is being arrived at in the latter series (a score of .36 rather than the .44 score mentioned above).

### *Reliability scores*

Reliability scores have been calculated at the end of the coding process, that is, once all the scenes had been coded by the researcher, to avoid any prejudices in the researcher's coding behavior. As said before, the scenes were coded by the researcher in random order. Moreover, the researcher did not know whether the scenes to be coded concerned an answer to the pre- or the posttest.

Regarding the overall reliability, that is the degree to which exactly the same cognitive maps (i.e. exactly the same concepts and relationships) were arrived at by different coders, both a stability and a reproducibility score have been calculated. However, one should take into account, that although it is interesting to know to which degree cognitive maps can be reproduced, it is more important to know to what agree the individual

[114] At the time the sample of eight scenes was coded, however, the dictionary by means of which the concepts were to be translated consisted only of 86 concepts. This is still large enough to not to worry about accidental agreement.

variables are affected by reliability tests. Hence, once an overview has been given of the overall reliability of the cognitive mapping procedures, a more detailed description of the degree to which scores on each of the variables based on the cognitive maps, have proven to be stable and reproducible.

Regarding the overall stability of the coding process (see figure 6.5, the fourth column, with the A1A2 on top representing the stability of the three coders), that is the degree to which the very same coder can reproduce the very same overall outcomes (i.e. cognitive map), an average stability score of 76 per cent was found with respect to the kind (not the number) of relationships identified from the texts. The stability of the kind (not the number) of concepts identified from the written texts, however, was found to be equal to 83 per cent. The reason why we deliberately say 'kind' of concepts and relationships rather than 'number', is first of all because only unique concepts and relationships are included in cognitive maps; the number of times these concepts and relationships are used is not taken into account. Moreover, the comparisons between and within coders are based on the kind of concepts and relationships arrived at and not on the number of concepts and relationships included in one's cognitive map. Consequently, coders who have coded exactly the same number of concepts and relationships, but have no concepts and relationships in common, will be given a reliability score of zero, at the assessment of the process' overall reliability.

As far as the overall reproducibility is concerned (depicted in figure 6.5, the third column, with the capital X-average on top), scores were found of 64 per cent with respect to the number of identical relationships and 77 per cent for the number of identical concepts. According to Krippendorff (1980), reliability scores in between 65 and 80 per cent, allow only for tentative conclusions, whereas scores higher than 80 per cent are needed to be able to be in the position to make more sound inferences. Applying these criteria to the reliability scores found in the present study, some concerns exist for obviously the reproducibility score is far from meeting the 80 per cent criteria. The score on stability however, does approach the 80 per cent level, as a result of which it can be claimed that, as long as it can be made clear that the procedure followed by the researcher is a valid one, the cognitive maps derived at by the coding procedure can be used to make scientifically acceptable inferences.

However, when we look at the reproducibility scores more carefully, we see that the low reliability between coder B and the other two coders (A and C) in particular, can be held responsible for the relatively low overall reproducibility score. As shown in figure 6.5, the overall correspondence between coders A and B (cf. column AB) is equal to 73 and 58 per cent and an overall correspondence between coders B and C (column BC) is found of 74 and 56 per cent. The reproducibility score between coders A and C (column AC) however, is as high as 84 and 74 per cent, which is much more acceptable than the average reproducibility score of 77 and 64 per cent that was presented at the beginning of the presentation of the reproducibility scores. Factors that may account for the differences between coders A, C and coder B are differences in the degree to which the coders were committed to the coding process and felt engaged in the success or failure of the research, the degree to which one was familiar with the material to be coded, the *object of research* (Muskens, 1980, p. 218), and the degree to which the coders understood the procedures to be followed to code the material (Muskens, 1980).

As mentioned above, comparing the cognitive maps with respect to their ***content***, that is, the relationships and concepts included in it, is the most rigorous reliability test that can be carried out, for many of the indicators that are based on the cognitive map, are not based on the exact content of the map at all. For instance, if we look at the indicator of endogeneity, the number of concepts and relationships labelled 'endogenous' is of importance, rather than the exact content of the concepts and relationships. In other words, a cognitive map consisting of two concepts and one relationship may be extremely unreliably coded if two completely different concepts are arrived at by the two coders, but as long as the concepts identified have the same value with respect to endogeneity, the outcomes of the coding process in terms of the indicator, will be the same. The reliability scores regarding the exact indicator thus are likely to be much higher than the overall reliability scores presented above. In figure 6.5, an overview is given of both the overall reliability scores and the reliability scores for each of the individual indicators used in the study:

Figure 6.5: Stability and reproducibility scores: overall values and per indicator

| | *reliability scores* | | | | |
|---|---|---|---|---|---|
| indicator | reproducibility | | | | stability |
| A, B, C refer to 3 coders | AB | AC | BC | $\bar{X}$ | A1A2 |
| *overall score on reliability* | | | | | |
| identical concepts | 73 | 83 | 74 | 77 | 83 |
| identical relationships | 58 | 74 | 56 | 63 | 76 |
| *reliability scores per indicator* | | | | | |
| endogenous concepts | 76 | 87 | 77 | 80 | 96 |
| endogenous relationships | 53 | 67 | 50 | 57 | 78 |
| exogenous concepts | 80 | 94 | 74 | 83 | 67 |
| exogenous relationships | 50 | 75 | 41 | 55 | 66 |
| connectivity | 87 | 92 | 94 | 91 | 91 |
| chaining | 92 | 89 | 87 | 89 | 93 |
| # of feedbackloops | 88 | 88 | 100 | 92 | 88 |
| # of concepts from external model | 90 | 89 | 86 | 88 | 97 |
| # of relationships from external model | 62 | 68 | 63 | 65 | 62 |
| *average of individual indicators* | 75 | 83 | 75 | 78 | 82 |

It is shown in figure 6.5, that much higher reliability scores can be obtained if one focuses on the individual indicators. Notice that to arrive at reliability scores for the individual indicators, a somewhat different procedure was followed because the content of the observations no longer could be used to decide whether two observations were identical. Hence, to determine the agreement among the coders, it was decided to determine the percentage of differences between the coders first, and subtract this percentage from unity, to arrive at an agreement score. Obviously, if no differences between the coders are found, the percentage of differences is equal to zero, as a result of which a reliability score of unity is being arrived at. To illustrate this procedure, look at the following example. Say, three scenes have been coded twice by the very same coder. The first time the scenes were coded, eight, seven, and five concepts were found (coded). The second time the coding process was carried out, however, seven, seven, and six concepts were being arrived at. To determine the stability score belonging to this example, the differences between the first and second time the coding was carried out have to be identified first. In the present example, a difference score of two is found; two concepts were missed (or two concepts added, depending on which of the two maps is considered to be the most appropriate one). To determine the importance of these two concepts (how bad is it to have missed two concepts?), a relative score is being calculated. For this, we divide the number of differences (in our case: two), by the smallest number of observations (to be sure we have the most rigorous measure possible), as a result of which a 'difference score' of 2/20 or .10 is being arrived at. The agreement score, indicating the degree to which the coding process is capable of bringing about the same outcomes with respect to a particular indicator, hence is equal to 90 per cent.[115]

Based on figure 6.5, one can conclude that the pattern found in the overall reliability scores also exists in the scores for each of the individual reliability scores: scores regarding relationships between concepts are lower than scores that concern concepts. In other words, it seems easier to identify the concepts that make up a cognitive map, than the relationships between these concepts. Moreover, note that because only a small number of concepts qualified as exogenous (as a result of which the number of exogenous relationships is low as well), reliability scores on this indicator need to be considered with care for most of the scores are based on only one or two entries. So in case both coders code one exogenous concept, a 100 per cent reproducibility score is being arrived at, whereas if only one of them coded an exogenous concept, the reproducibility score immediately drops to 0 per cent.

In sum, although there is room for improvement with respect to the reliability of the cognitive mapping procedure in general, the figures presented in figure 6.5 show that, generally speaking quite acceptable overall stability scores can be arrived at, and that the reliability scores of the individual indicators are substantially higher than the ones found for the cognitive map as a whole, as a result of which it can be concluded that the cognitive map constructed in the present study can be used to draw conclusions from with respect to the domain-specific and strategic content of the way in which people look upon a problem.

However, it should be said that, because of, among others, the tedious and laborious character of the coding process, alternative means to arrive at cognitive maps should be looked for seriously. In the concluding chapter, we will come back to this issue and suggest another way to arrive at cognitive maps without having to tranfer written text in cause-effect relationships, as a result of which reliability no longer will be an issue, and outcomes of the evaluation can be arrived at at an earlier stage.

## 6.3 CONSTRUCTION OF THE VARIABLES

Having discussed the construction of the cognitive maps by means of which values can be arrived at for each of the indicators presented in figure 4.13, account must be given to the construction of the variables that are made up of more than one indicator. The description of the way in which the variables have been constructed will focus on the potentially specifying variables to start with - an overview will be given of the way in which variables concerning time investments, background variables, evaluation of the program in terms of components and aspects, and the evaluation of the effects of the program, have been constructed on the basis of their individual indicators. Following the discussion of how the potentially specifying variables have been constructed, the construction of the dependent variables will be focused upon, that is, an overview will be given of the way in which the indicators concerning strategic and domain-specific knowledge have been combined, to arrive at a measure for the variables that are of importance to measure the theoretical concepts outlined in Chapter Three.

[115]This procedure follows closely the procedure taken by Axelrod (1976), and Vennix (1990).

### 6.3.1 CONSTRUCTION OF THE POTENTIALLY SPECIFYING VARIABLES

As said above, the description of the construction of the potentially specifying variables will subsequently focus on the amount of time participants have spent on the participative policy modelling program, the way in which they look upon the program in terms of components and aspects, and the effects participants believe have been brought about by the program in which they participated.

To measure the amount of time spent on the program, the participants were asked to state how many hours they had spent preparing for the sessions, and it was measured how many sessions they had attended. Since the number of hours taken by each of the sessions is known, a grand total for time investment can be arrived at by adding the scores on both indicators.

In order to qualify as a subject in the present study, participants were required to have filled out the pre- and posttest. As a consequence, of the 29 persons who did take part in the program, only 26 could be considered as subjects, that is, only 26 had completed both pre- and posttest. Of the remaining three participants, two had only completed the posttest, whereas one had only filled out the pretest. Of the 26 participants who had completed both pre- and posttest, 4 had not answered the questions seriously on either the pre- or the posttest. To assess whether a person had tried to answer the questions in a serious manner, attention was paid to the number of lines that were written down to give an answer. If to any of the three questions an answer was given that was only one line long (recall, the participants were asked to state what they thought the effects of a particular policy measure would be *and* to explain why they believed the effects would be what they thought it would be), the answers given to either pre- or posttest were considered to be an invalid representation of the person's knowledge about the problem (and solution) at hand, as a result of which the participant was removed from the list of subjects.

The group of subjects that did meet the criterion of having seriously completed the pre- and posttest, is called the *intention to treat group*. They are called the *intention* to treat group because being part of this group not necessarily means that they did take part in the program in such a way that effects can be expected. In the present study, the intention to treat group consists of twenty-two persons.

Obviously, having attended only one of the four sessions is not likely to bring about a change in one's conceptualization in a so complicated matter as the Dutch Health Care system. Not only because of the complexity of the system under consideration, but also, as Axelrod (1976, p. 230) states *"...important segments of a cognitive map can remain stable over relatively long periods of time"*. Cognitive maps (i.e. conceptualizations) do not change easily, as a result of which it is not likely that incomplete treatment is able to bring about the desired effects (changes) in conceptualization. Hence to examine the effects of participation in participative policy modelling, those participants who really did take part in the program are taken together and considered as the *on-treatment group*.

To distinguish between those who did and those who did not really participate in the program, use was made of the scores concerning the amount of time spent on the program. Participants who did not prepare for three hours or more in total for the first three sessions (the fourth session did not require any preparation), <u>and</u> who did not attend two of the first three sessions (the first three sessions are considered to be the most important ones), failed to qualify for the on-treatment group. Since five participants failed to meet these criteria, one of which had already dropped out due to failing to meet the number of lines requirement, 18 participants in total managed to qualify for the on-treatment group. Whether any substantial differences exist between the intention to treat and the on-treatment groups exists with respect to, for instance, background characteristics or participant-based evaluation of the program, will be discussed in the next chapter concerning the outcomes of the present study.

### *Background variables*

With respect to the background variables included in the questionnaire, note that most of the variables have already been presented in figure 5.3 to illustrate the differences between and within the two departments participating in the present study. Moreover, since no complex computations have to be carried out to arrive at values for these variables, no account needs to be given of a construction process. In the summary of the present chapter, an overview will be given of all the variables, including the background variables, used in the present study to answer the research questions presented in Chapter Three.

### *Evaluation of the program*

As already indicated in Chapter Four, it was decided to make use of a priori scales to arrive at scores for the three components that make up the participative policy modelling program and the five aspects that can be distinguished as well. The reason why it was decided to make use of a priori scales to measure a variable such as 'the degree to which participants felt positive about a component of the program' (e.g. the workbook) rather than following Vennix (1990) and carry out a factor analysis on a 'flattened' data matrix, is because no correlation between the items that make up some of our variables was expected, despite the fact that it was believed that some of the items had to be taken together to arrive at a value for particular variables. To illustrate, to arrive at one value for the appreciation of, say, the workbook component, it was found that indicators such as the degree to which the workbook had been interesting, useful, easy, and somewhat long, had to be taken together, for all contribute to the degree to which such a component is appreciated. However, these individual indicators do not necessarily correlate for they refer to different dimensions of one and the same complex variable.

However, since it is our objective to compare the present study's outcomes as much as possible to the outcomes of the study carried out by Vennix (1990), it was decided to follow the process carried out by Vennix to arrive at scores for each of the combined variables as well. As a result, Factor Analyses were carried out per session (that is, based on the items that concern one and the same session), on averages of sessions (that is, by averaging the scores on each of the items over all four sessions), per dimension (or aspect) per session (that is, by averaging all items per session that are related to one dimension or aspect such as usefulness), and by averaging per session and then over sessions (e.g. by first calculating the average value of, say, usefulness over all three components, and subsequently determining the average of the averages of each of the four sessions), to see whether any underlying dimensions could be found, and secondly, to examine in case underlying dimensions were found, if they were identical to the ones arrived at in the Vennix study. As expected, none of the analyses carried out resulted in interpretable dimensions; no satisfying factor solutions could be arrived at.

Based on this, it was decided to use the a priori scales outlined in Chapter Four to arrive at values for each of the evaluation variables (i.e. the aspect variables and component variables). Note that to arrive at scores for each of the components and aspects, all indicators referring to a particular aspect or component will be averaged. Since all indicators use 5-point scales (scores ranging from 1 to 5), the scores for each of the indicators cary from 1 (the lowest extreme) to 5 (the highest, most positive extreme). For an overview of the variables thus obtained, the reader is referred to the summary of the present chapter.

### *Evaluation of the effects of participative policy modelling*

Recall that the participants were not only asked to express their views on the program in which they had been participating in terms of aspects and components, but that they were also given the opportunity to state whether the program had succeeded in bringing about any changes in their own conceptualization of the problem being focused upon. For this, they were asked to answer questions number 26 to 35 at the end of the third session. With respect to the variables that have been measured by more than one item (in figure 4.6 and 4.7 an overview has been given of the items included in the questionnaire to assess a change in domain-specific and strategic knowledge), average scores have been calculated to arrive at values for these variables. For example, to determine the value of a variable such as 'awareness of others' point of view', an average score based on items number 17, 24, 25, 33, 34, 42, and 43 is calculated. Regarding the variables that are made up of only one item, no computations are carried out. The average value of this item (averaged over all participants) is taken as a measure of the variable itself.

## 6.3.2 CONSTRUCTION OF THE DEPENDENT VARIABLES

With respect to the construction of the variables by means of which changes in domain-specific and strategic knowledge are being evaluated, several issues will have to be dealt with. First of all, account will be taken of the indicators that have been removed because of low reliability, low frequency, and/or other problems. Next, a description will be given of the way in which some of the variables that are based on more than one indicator, are constructed. Finally, some attention will be paid to the way in which it is intended to combine the subcriteria outlined in Chapter Three, to assess the effects of participative policy modelling on domain-specific and strategic knowledge rather than on separate variables only.

### *Elements to be included in the analyses*

In line with Vennix (1990), it was decided to remove the number of quantifications[116] from the list of variables to be included in the analyses because of its extremely low frequency. Since the number of quantifications had not been included to assess the degree to which people change their strategic and domain-specific knowledge, removing the variable from the analyses need not be considered as a dramatic change at all[117].

Moreover, the multi-disciplinarity indicator was excluded from the analyses because even at the pretest, only one concept succeeded in qualifying as a mono-disciplinary one. The extremely low variance on this variable made us decide to drop the variable from the analyses. In the concluding chapter, some attention will be paid to the absence of significant differences in the use of concepts (and relationships) between the two departments paricipating to the present study, despite the clear difference in the nature of their activities and the educational background of the participants.

As far as the reliability of the indicators is concerned (cf. figure 6.5), the relatively low scores on reproducibility and stability regarding the number of exogenous relationships clearly suggest that some caution should be taken in the use of this individual indicator as a basis for inferences regarding the domain-specific knowledge included in a person's conceptualization. As far as the other indicators are concerned, it is felt that their reliability scores do legitimate their use in the present study to a full extent.

To check whether any of the indicators had to be removed due to skewness, Kolmogorov-Smirnov Goodness of fit tests were carried out on the gain-scores (posttest minus pretest) that are used to answer the research questions outlined in Chapter Three.

---

[116]The quantifications that are neither included in the Vennix (1990) study nor in the present study concern the relationships between two variables (e.g. an increase in X will result in a *10* per cent increase in Y).

[117]The number of quantifications had been included to allow for comparisons with the Vennix (1990) study. It is interesting to note that in both the present study and the study carried out by Vennix, hardly any of the causal relationships found in the written texts were quantified. As a consequence, in both the present studies and the study carried out by Vennix (1990), it was decided to remove the number of quantifications as a variable from the list of variables to be examined in more detail.

The results of this test showed that none of the remaining 10 indicators was skewed as a consequence of which none of the indicators had to be excluded from the analyses.

## *Combining indicators: variables*

Since not all the aspects of domain-specific and strategic knowledge are made up of one indicator (cf. figure 4.13), account must be taken of the exact way in which the indicators have been combined to arrive at values of the variable. To start with, a presentation will be given of the way in which the variables that belong to the domain of domain-specific knowledge have been constructed. This will be followed by an overview of the construction of the variables that make up the strategic knowledge dimension of the present study.

Regarding the aspect of endogeneity, a distinction is made between the number of endogenous concepts and the number of endogenous relationships. To arrive at a score on endogeneity, the number of exogenous concepts is substracted from the total number of concepts of either the pre- or the posttest. The same is done with respect to the number of endogenous relationships: to obtain a value for the number of endogenous relationships, the number of exogenous relationships is substracted from the total number of relationships, for relationships (and concepts as well) can only be endogenous or exogenous (note that the 'or' is an exclusive 'or'). Moreover, since the number of concepts and number of relationships are highly correlated ($r=.87$, $p<.01$ for the pretest, and $r=.89$, $p<.01$ for the posttest), it was decided to add the two indicators, to arrive at a grand total for the cognitive map's endogeneity, representing the number of entries in that domain. The same procedure was followed with respect to the exogenous concepts and relationships included in the cognitive map, resulting in one exogeneity score, and the number of concepts and relationships used from the external model, resulting in one external model score.

To arrive at a value for exogeneity, the number of exogenous concepts and the number of exogenous relationships have to be determined first. In order to arrive at values for both indicators however, the number of concepts and relationships from a field that is different from the field that is being referred to by the pre- or posttest question, needs to be added to the number of concepts and relationships that are exogenous in the second sense of the word, in that they affect the health care system but are themselves not affected by the system.

Note that the variables referring to the degree to which the external model has been incorporated in a participant's conceptualization, can be calculated directly on the basis of the indicators measured by means of the cognitive map. No mathematical computations have to be carried out to determine the value of 'external model concepts' and 'external model relationships'. However, as mentioned above, the two indicators are added to arrive at a grand total for the variable called 'external model'.

The strategic knowledge variables 'feedbackloops' and 'time-phase relationship' need not be explained in more detail because their values follow their measurements directly.

The 'chaining' variable by constrast, is constructed in a somewhat more complex manner. Its value is based on the values of the following four indicators: the number of feedbackloops, the length of feedbackloops, the number of maximal paths and the length of maximal paths. By calculating the average length of the maximal paths and feedbackloops incorporated in the answers to the pre- and posttests, a value for the 'chaining' aspect can be arrived at.

Finally, to determine the value of 'connectivity', the total number of relationships is divided by the total number of concepts.

In summary, the variables included in the present study are the following:

Figure 6.6: Variables included in the present study

*Dependent variables*

| main criterion | based on cognitive map | based on the questionnaire |
|---|---|---|
| domain-specific (II and IV) | endogeneity (concepts & relationships) | domain-specific knowledge in general |
| | exogeneity (concepts & relationships) | endogenous and exogenous knowledge |
| | external model (concepts & relationships) | |
| | | inter-individual change in domain specific knowledge |
| strategic (I and III) | feedbackloop | feedbackloops |
| | chaining | chaining |
| | connectivity | connectivity |
| | time-phase relationship | |
| awareness of others' point of view (V) | | awareness |

*Potentially specifying variables*

*background variables*

| | | | |
|---|---|---|---|
| age | educational background | department | organization |
| area of expertise | years with the firm | policy experience | gender |
| time-investment | hours of preparation | hours attending | |

*evaluation of the program*

| components | aspects | |
|---|---|---|
| workbook | usefulness | interesting |
| small-group activities | easiness | time spent |
| plenary session | learned from it | |

## 6.4 SUMMARY

In the present chapter, the procedures followed to transform the pre- and posttest into data that can be analyzed have been focused upon. To start with, an overview was given of the cognitive mapping approach by means of which the texts were transformed into cognitive maps that consist of concepts and relationships. Following this overview, the validity and reliability of these procedures were discussed. Finally, the construction of the variables that play a major role in the present study's evaluation (dependent variables and potentially specifying variables) were examined. This resulted in an overview of the variables included in the present study and the way in which they have been operationalized.

# CHAPTER 7: RESULTS

## 7.1 INTRODUCTION

In this chapter, an overview will be given of the outcomes of the present study. The outcomes presented in this chapter, serve to answer the major research questions outlined in Chapter Three. The empirical results of the research process carried out to answer these questions will follow the order in which they have been presented in Chapter Three and Four. As a consequence, the presentation of the outcomes of the study, that is, the changes brought about by the participative policy modelling program in the knowledge people have about a particular policy problem, and the knowledge they have about the way in which other participants look upon the problem, will focus on the *individual* changes brought about by the program first (research questions I, IIa, and IIb). This will be followed by a description of the *inter-individual* changes the participative policy modelling method is able to bring about (research questions III, IVa, and IVb). It is only when the changes related to the knowledge people have about a the policy problem at hand have been presented, that the outcomes related to the knowledge and/or *awareness* participants have about each other will be focused upon (research question V).

However, prior to assessing the outcomes of the present study in terms of the research questions to be answered, a brief discussion of the characteristics of the cognitive maps derived from the text written by the participants will be given. This to provide some (empirical) understanding of the data that have been arrived at, for it is our belief that such a brief description of the empirical material will contribute to one's understanding of the outcomes. The description of the general characteristics of the cognitive maps will focus on features such as (change in) size (mean and variance), number of new concepts and relationships, and number of concepts and relationships.

In the description of the outcomes, use will be made of the distinction between the 'intention to treat' and 'on-treatment' groups. Recall that the former consists of all the people who filled out a pre-and posttest seriously, whereas the latter consists of people who can truly be called participants because of the fact that they attended most of the sessions and spent a substantial amount of time preparing the sessions.

Once the main research questions have been answered for both the intention to treat and on-treatment group, the outcomes of additional analyses (e.g. analyses in which the effects of modifying or specifying variables are being examined) will be presented for the on-treatment group only. The reason for this is that first of all, to examine the effects of participative policy modelling one should be concerned with people who indeed have been participating in the program. As a consequence, it makes sense to present the effects of potentially specifying variables on the relationship between treatment and dependent variables only for those who have indeed been subjected to the treatment. Secondly, as will become clear in the description of the major research questions, the differences between the two groups, regarding the way in which their participants changed their conceptualization, is limited to such an extent that describing the outcomes for both groups would be somewhat repetitive[118].

Prior to the description of the effects brought about by the potentially specifying variables, account will be given to the values that have been found for each of these potentially specifying variables, that is, the background variables, the amount of time spent on the program, and the participant-based evaluation of the program. Not only because a description of these variables is needed to arrive at a clear understanding of

[118] Since the people belonging to the on-treatment group are all part of the intention to treat group as well, and only four of the twenty-two subjects of the intention to treat group are not part of the on-treatment group, one can expect not to find too many differences between the on-treatment and intention to treat group.

these additionally specifying analyses, but also because knowledge of the values of these potentially specifying variables contributes to our understanding of the application of the participative policy modelling method in the present study. Moreover, based on the description of the participant-based evaluation of the program, some conclusions with respect to the participative policy method itself will be drawn in Chapter 8, that is, some recommendations will be given as to where there is room for improvement, and how to improve the participative policy modelling method.

Finally, once a description has been given of the analyses carried out on the basis of the research carried out in the present study, comparisons with the Vennix study will be made, because it is most close to the present study as a result of which additional knowledge can be acquired by comparing the outcomes of both studies. Although a substantial correspondence exists between the Vennix (1990) study and the present one, some adjustments will have to be made to the operationalization of the present study's variables to be able to compare the outcomes of both studies. The number of lines of the texts written by the participants, for instance, will have to be taken into account in order to be in the position to compare the present study with Vennix (1990).

## 7.2 GENERAL DESCRIPTION OF THE COGNITIVE MAPS

In the present study, the participants were asked to answer three questions before and after completing the participative policy modelling program. To acquire some understanding of what the cognitive maps derived from these answers look like, figure 7.1 has been included:

Figure 7.1: General characteristics of the cognitive maps

| | intention to treat (N=22) | | on-treatment (N=18) | |
|---|---|---|---|---|
| | pretest | posttest | pretest | posttest |
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) | $\bar{x}$ (sd) | $\bar{x}$ (sd) |
| # of lines | 19.82 ( 7.48) | 16.05 (5.19) | 19.22 ( 7.48) | 16.61 (4.68) |
| # of concepts | 24.46 (10.34) | 22.05 (6.99) | 24.78 (11.24) | 22.56 (7.03) |
| # of relationships | 26.50 (13.85) | 22.55 (8.91) | 26.72 (14.84) | 23.44 (9.02) |
| # of new concepts | | 11.82 (6.69) | | 12.28 (6.99) |
| # of new relationships | | 18.23 (9.28) | | 19.00 (9.44) |

As shown in figure 7.1, the average number of lines written down on the pretest is equal to about 20 whereas on the posttest, on average 16 lines were written down (this goes for both the intention to treat and on-treatment group). It is our believe that one of the most important reasons why a decrease in the average number of lines is found, is because the participants on average spent less time answering the questions of the posttest than they did on the pretest[119]. This decrease in average number of lines is reflected in a decrease in the average number of concepts and relationships[120]. On average, the cognitive maps

[119]Note that the participants were not asked how much time they had spent answering the questions. However, as illustrated in Appendix 12, it is not always difficult to tell on which of the two measures more time had been spent. Generally speaking, it was found to be much more difficult to have the participants fill out the posttest than it was to have them complete the pretest.

[120]The number of lines and the number of concepts and relationships do not necessarily correlate, because the number of concepts and relationships concern the number of *unique* concepts and relationships. The reason why it was decided not to divide all dependent variables by the number of lines to arrive at variables

of the posttest, representing how the participants look upon the problem having participated to the program, are smaller in size than the cognitive maps that are based on the pretest. Although this decrease in the number of concepts and relationships can be regarded as an improvement in that the program may have made the participants aware of the domain-specific elements that are really important in considering the problem, and/or has helped them to arrive at a more condense representation of the problem (e.g. an increase in connectivity), it seems more likely at this stage, because of the limited amount of knowledge people tend to have with respect to ill-structured problems, that these statistics do suggest that some decrease in the knowledge people have about the problem at hand is being brought about by the program.

Another interesting feature of the cognitive maps concerns the number of concepts and relationships participants use in the posttest that were not used in the pretest, i.e. the number of *new* concepts and *new* relationships. It is interesting to see that although the size of the cognitive map (the number of concepts and relationships used to refer to the problem) does not change substantially (it changes from 24 concepts on the pretest to 22 concepts on the posttest, and from 26 relationships on the pretest to about 22.5 relationships on the posttest, for both the intention to treat group and the on-treatment group), people do replace a substantial part of their conceptualization in, say, six months time (the time between pre- and posttest). In figure 7.1, it is shown that of the 22 concepts used on average on the posttest, 12 are new so that only 10 concepts are used on both pre and posttest. The very same holds for the relationships used in the pre- and posttest: of the 26 relationships that appear in the answers given to the pretest, only 4 are used in the posttest. The additional 18 relationships that are used to describe the problem can be considered as new relationships, in that they were not used by the participants on the pretest. In terms of percentages, these figures show that just more than 50 per cent of the concepts and 78 per cent of the relationships that are used on the posttest can be regarded as new. Based on these figures, one can conclude that people do change the concepts and relationships they use to describe the problem at hand dramatically[121]. As to whether the kind of change in conceptualization is in line to the changes expected to take place (research questions I to IV), will be examined in more detail in the sections to come.

---

such as 'the number of exogenous concepts per line', and 'the number of endogenous variables per line' is because we are primarily interested in the way in which a person is looking upon a particular problem, as expressed in a specific discourse, rather than in the way in which a person expresses him- or herself, that is, the writing style a person has, which probably is even more difficult to change than one's conceptualization. As will be elaborated upon later on in the present chapter, students participating in the Vennix study, for instance, needed almost twice as much lines to express the very same number of concepts and relationships than the participants in the present study did. So, although the amount of domain-specific knowledge expressed in the texts is almost identical for both groups, they differ substantially in the number of lines that are needed to express this knowledge. Another reason why the present study did not follow Vennix (1990, p. 152) by dividing all variables by the number of lines written down by the participants, is because we do not believe that because a person writes down more lines, (s)he will automatically have used more concepts and relationships. Not only because we are dealing with the number of unique concepts and relationships as explained before, but also because it is our belief that people who use more lines do so *because* they use more concepts and relationships (not necessarily new ones). Since we assume the causality to be of a different direction, we need not be worried about a confounding effect of the number of lines used to frame an answer.

121 Due to the fact that no control group is used in the present study, it is difficult to ascribe this change to the participative policy modelling method. On the other hand, one should take into account that mental maps are said to be relatively stable, they do not change much over time (Bonham & Shapiro, 1976; Axelrod, 1976).

## 7.3 MAJOR RESEARCH QUESTIONS

Having given an overview of the content and size of the cognitive maps based on the pre- and posttests filled out by the participants, account should be taken of the five main research questions that are focused upon in the present study. The answers to the first four research questions are based on information that stems from the pre- and posttest completed by the participants, and the questionnaire filled out at the end of each session. The fifth question however, will be answered on the basis of the questionnaire only, for we have no indicators that are based on the cognitive mapping approach by means of which a change in awareness of the other participants' point of view can be measured.[122]

Regarding the first four research questions, a brief summary of the research question will be given first, resulting in a description of the hypotheses that need to be answered with respect to each of the main research questions. This will be followed by a presentation of the outcomes that are based on the cognitive map, and an overview of the results that come from the questionnaire filled out by the participants. Finally, the findings of both approaches (the cognitive mapping approach and the questionnaire approach) will be compared and conclusions will be drawn with respect to the question as to whether the participative policy modelling has been able to live up to its expectations, that is, bringing about the change it was expected to bring about. Note that the description of the outcomes will be given for both the intention to treat and the on-treatment group[123].

### 7.3.1 INDIVIDUAL STRATEGIC KNOWLEDGE (I)

#### *Assessing the effects on the basis of cognitive maps*

The first research question that needs to be considered concerns the degree to which participants have been able to change their strategic knowledge with respect to a particular problem, that is, the way in which they organize their domain-specific knowledge. The variables that were identified to measure such a change in strategic knowledge are the number of feedbackloops, the cognitive map's connectivity, chaining, and the number of time-indicators. The hypotheses that will be focused upon in the present section hence are the following:

*Hypothesis I-1: on average, the **feedbackloops**-score will be higher on the posttest than on the pretest*

*Hypothesis I-2: on average, the **connectivity**-score will be higher on the posttest than on the pretest*

*Hypothesis I-3: on average, the **chaining**-score will be higher on the posttest than on the pretest*

---

[122] Prior to discussing the outcomes, note that to control for a maturation effect, analyses were carried out to assess whether a correlation between age (and/or number of years with the firm) and pretest scores existed. Since no significant correlations were found (with the exception of the correlation between the variables 'age' and 'exogeneity score' ($r=.69$, $p< .01$) and 'number of years with the firm' and 'exogeneity score' ($r=.57$, $p< .05$), maturation cannot be regarded as a too great threat to the validity of the present study's outcomes.

[123] Note that in order to be in the position to attribute the effects found in this study to the participative policy modelling method rather than to the fact that people grow older (and wiser) at the time they take part in the sessions (i.e. 'maturation' as discussed in detail in Chapter Four), analyses were carried out to see whether any significant correlation existed between pre-test scores and variables such as 'age' and 'number of years with the firm'. Since, no significant correlations were found (older people did not have more knowledge either of a domain-specific nature or of a strategic nature, and people who had been with the firm for a longer period of time did not have higher pretest scores, the potential threat of maturation can be ruled out.

*Hypothesis I-4: on average, the **time-indication**-score will be higher on the posttest than on the pretest*

To test these four hypotheses, analyses of variance (ANOVA) were carried out. The results are the following:

Figure 7.2: Individual change in strategic knowledge

| Strategic | intention to treat | | | on-treatment | | |
|---|---|---|---|---|---|---|
| | pretest | posttest | | pretest | posttest | |
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) | p-value | $\bar{x}$ (sd) | $\bar{x}$ (sd) | p-value |
| feedbackloops | .46 ( .67) | .55 ( .86) | .61 | .50 ( .71) | .56 ( .86) | .79 |
| chaining | 6.46 (1.73) | 6.21 (1.99) | .37 | 6.60 (1.86) | 6.48 (2.07) | .79 |
| connectivity | 1.03 ( .15) | 1.00 ( .12) | .37 | 1.04 ( .15) | 1.02 ( .11) | .49 |
| time-indications | .68 (1.04) | .32 ( .57) | .09 * | .56 ( .92) | .33 ( .59) | .26 |
| N | 22 | 22 | | 18 | 18 | |

Concerning the variable 'number of feedbackloops', it is clear that some improvement has been achieved (a 20 per cent increase at the intention to treat group and an 11 per cent increase at the on-treatment group). However, because the standard deviations are relatively large, neither of the two differences are significant.

At the pretest, participants include on average .46 (intention to treat) and .50 (on-treatment) feedbackloops in their answers. In other words, only one out of every two participants has a feedbackloop included in his or her conceptualization; policy makers apparently have great difficulty in thinking in terms of feedbackloops. The results are in line with Axelrod (1976), Weick & Bougon (1986), and Hall, Aitchison, Kocay, and Li (1989), who state that when *"...cognitive maps are coded from documents, it is rare to find loops"* (Weick & Bougon, 1986, p. 121-122). However, they do not correspond to the Vennix study in which an average number of feedbackloops of 9.21 on the pretest and an average number of loops of 13.35 on the posttest is being found (Vennix, 1990, p. 192). A possible explanation for the relatively large number of feedbackloops found in the Vennix study is that the participants in that study were asked to answer a question about an economical problem, a problem about which a substantial amount of knowledge is available, whereas in the present study participants were asked to conceptualize an ill-structured problem about which hardly any available, ready-to-use knowledge existed. Another explanation for this difference is that econometric models contain more relationships than system dynamics models do - econometric models generally speaking have a higher connectivity or density than system dynamics models have - as a result of which artificial (unintended) feedbackloops are much more likely to be included in the total number of feedbackloops. In the Vennix study (1990, p. 190), an increase in the density of the cognitive maps and the total number of relationships was found, as a result of which indeed the likelihood of artificial feedbackloops increases.

As far as the connectivity of the cognitive map is concerned, values of about unity have been found on the pre- and posttest, for both the intention to treat and on-treatment groups. Although the change in connectivity can be considered as insignificant, some indication exists that the participants' conceptualizations have become slightly less integrated, despite the fact that they use less concepts.

Regarding the 'chaining' variable, representing the dynamics that can be found in the conceptualization in terms of the average length of the cause-effect chains, a slight decrease in the average length of a chain is found for both the intention to treat and on-treatment groups. Although these changes obviously are in the opposite direction of what was expected, they are so small that they have to be considered as being not significant.

Finally, in contrast to our expectations, a decrease in the average number of time-indications was found for both the intention to treat and the on-treatment group. The intention to treat group decreased its score with about 50 per cent, whereas a decrease of about 40 per cent was found for the on-treatment group. Again, because of the relatively small size of the number of subjects and the relatively large standard deviations, only the first of the two differences, concerning the intention to treat group, proved to be significant.

Summarizing the effects brought about by the participative policy modelling method assessed by the answers given to pre- and posttest, it is clear that no substantial improvement can be concluded to on the basis of the statistics presented in figure 7.2. As a matter of fact, on three of the four variables a small decrease rather than an increase is found, although none of these changes can be considered as significant. As far as the differences between the intention to treat and on-treatment groups are concerned, differences are small in size, but show that on most of the variables the on-treatment group outperforms the intention to treat one.

### *Assessing the effects on the basis of the questionnaire*

With respect to individual change in strategic knowledge brought about by the participative policy modelling method, participants were also asked to answer to some of the questions of the questionnaire (questions number 26, 27 and 29)[124]. Their answers given to these questions are the following (a five-point scale was used to assess how the participants felt about the strategic changes that had been brought about by the participative policy modelling program):

Figure 7.3: Strategic knowledge (questionnaire-based)

| | intention to treat | on-treatment |
|---|---|---|
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) |
| feedbackloops | 2.46 (.60) | 2.44 (.62) |
| chaining | 2.27 (.77) | 2.28 (.75) |
| connectivity | 2.36 (1.2) | 2.39 (1.2) |
| N | 22 | 18 |

As shown in figure 7.3, the people participating in the participative policy modelling method believe that on average the participative policy modelling method has contributed only slightly to their strategic knowledge. They seem, to a limited extent, of the opinion that feedbackloops are important to conceptualize the dynamics of ill-structured problems, however, this is not to say that they actually use or intend to use feedbackloops, as is illustrated by the extremely low number of feedbackloops found in their pre- and posttests.

Moreover, the relatively modest scores on chaining and connectivity, indicate that the participants are of the opinion that only a minor improvement has been made with respect to the degree to which the conceptualization is integrated, and the degree to which use has been made of cause-effect chains. As such, the participants do not feel a major change in the strategic knowledge by means of which they organize their domain-specific knowledge, has been brought about by the participative policy modelling program. Note that no major differences were found between the intention to treat and

[124]For a detailed description of the items included in the questionnaire, see Appendices 1 to 4.

on-treatment group; in both groups, the occurrence of a substantial improvement on the strategic knowledge dimension is being denied.

As a consequence, from a general point of view, the results based on the cognitive maps coincide with the outcomes that stem from the questionnaire, in that both show that no substantial change in strategic knowledge is found on an individual level. It is because of this lack of change that the hypotheses stated at the beginning of this section (I-1 to I-4), have to be rejected.

Having answered the first research question, the next section will be concerned with individual changes in *domain-specific* knowledge. As such, attention will be paid to the research questions IIa and IIb.

## 7.3.2 INDIVIDUAL DOMAIN-SPECIFIC KNOWLEDGE (IIa,IIb)

### *Assessing the effects on the basis of cognitive maps*

In this section, the question as to whether the program has succeeded in bringing about a change in the content of the participants' conceptualization of the problem will be focused upon. In order to prove that indeed a change in content has been brought about (this may well be a qualitative rather than a quantitative change: people need not use more concepts and relationships, but they may change the kind of concepts and relationships that are used to refer to the problem), use will be made of the variables 'endogeneity', 'exogeneity', and 'external model'. Note that the first two variables refer to research question IIa, whereas the 'external model' variable is referring to research question IIb (cf. figure 3.9 in Chapter Three). The hypotheses that can be derived from these research questions are the following:

*Hypothesis IIa-1:on average, participants have a higher* ***endogeneity****-score on the posttest than they have on the pretest*

*Hypothesis IIa-2: on average, participants have a higher* ***exogeneity****-score on the posttest than they have on the pretest*

*Hypothesis IIb-1: on average, participants have a higher score on* ***external model*** *variable on the posttest than they have on the pretest*

Regarding the change that was brought about in the domain-specific knowledge participants have about the problem(s) at hand, analyses of variance (ANOVA) were carried out to start with, to examine the differences found on each of the individual variables. The results of these analyses can be presented as follows:

Figure 7.4: Individual change in domain-specific knowledge

| Domain-specific | intention to treat | | | on-treatment | | |
|---|---|---|---|---|---|---|
| | pretest | posttest | | pretest | posttest | |
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) | p-value | $\bar{x}$ (sd) | $\bar{x}$ (sd) | p-value |
| endogeneity | 43.95 (18.66) | 37.82 (12.83) | .15 | 45.16 (19.43) | 39.11 (13.62) | .23 |
| exogeneity | 6.50 ( 6.55) | 6.77 ( 9.11) | .90 | 6.33 ( 7.19) | 6.89 ( 9.24) | .83 |
| external model | 24.36 ( 7.48) | 25.05 ( 9.16) | .73 | 24.00 ( 8.02) | 25.70 ( 9.81) | .46 |
| N | 22 | 22 | | 18 | 18 | |

With respect to the number of *endogenous* concepts and relationships (i.e. the variable 'endogeneity'), a difference in average score between the pre- and posttest is found for both the intention to treat and on-treatment group. However, the direction of the change is opposite to what was expected. In both groups the number of endogenous entries (that is concepts and relationships) is decreased by 6. Because the number of endogenous entries on the pretest is equal to 43.95 and 45.16, this decrease is equivalent to about 13 per cent. Due to the relative small number of subjects and substantial variance among these subjects, however, no significant change in endogeneity can be concluded to (p=.15 for the intention to treat group and p=.23 for the on-treatment group).

Regarding the *exogeneity* scores, only a slight difference between the pre- and posttest is found. In the 'intention to treat' group an increase in the average number of exogenous concepts and relationships is found of .27 (which is equal to approximately 4 per cent), whereas the 'on-treatment' group manages to increase its exogeneity score by .56, which is equal to about 8.5 per cent. However, although the change in both groups is in the right direction, it cannot be considered as a significant change (p=.90 for the intention to treat group and p=.83 for the on-treatment group).

It is interesting to note that the number of endogenous concepts and relationships is substantially larger than the number of exogenous concepts and relationships (a 7 to 1 ratio is found), indicating that indeed, as could be expected, people have a strong tendency to focus on concepts and relationships closely related to the problem, that is, concepts and relationships that come from the very same field (e.g. the subsystem of the general practitioner, or the subsystem of the medical specialist) rather than from one of the subsystems surrounding the subsystem the question was referring to.

The last domain-specific variable that needs to be taken into account concerns the degree to which a conceptualization corresponds to the *external model* created by and approved of by the participants. In figure 7.4, it is shown that in both the intention to treat and on-treatment groups an increase in the average number of concepts and relationships used from the external model is found. On the pretest, participants use on average 24 concepts and relationships from the external model, whereas on the posttest, 25 to 26 corresponding entries are found. Although this change is in the order of 7 per cent, their p-values show that none of them can be considered as significant (a p-value of .73 for the intention to treat group and a p-value of .46 for the on-treatment group is found). The number of elements that the preliminary model and the participants' conceptualizations have in common is substantial. On the pretest, 24 of the 50 entries correspond to the external model (about 48 per cent), whereas on the posttest 25 of the 45 entries qualify as in accordance with the external model (approximately 56 per cent)[125].

Moreover, it should be noted that with respect to all three variables, the on-treatment group outperforms the intention to treat group. Not surprisingly because the on-treatment group consists of people who have been participating whereas the intention to treat group is made up of participants who simply completed both pre- and posttest 'seriously'. On each of the three variables, the on-treatment group is found to score higher.

Summarizing the findings with respect to the individual changes in domain-specific knowledge based on the cognitive map constructed on the basis of the texts written to answer the pre- and posttest, the data suggest that a decrease in domain-specific knowledge is being brought about by the participative policy modelling method since the total number of entries (concepts and relationships) has been decreased. However, despite

[125] In the Vennix (1990) study, participants used even more concepts and relationships from the external model. It is believed that this is mainly due to the fact that the participants in the Vennix study were asked to solve a problem for which a considerable amount of knowledge (i.e. an accepted model of the problem) already existed. Hence the participants in the Vennix study were asked to solve a difficult but moderately-structured to well-structured problem. A more detailed description of the outcomes of the Vennix (1990) study in relationship to the present study's outcomes will be given in section 7.6.

this decrease in the total number of concepts and relationships, an increase in the number of exogenous concepts and relationships and the number of concepts and relationships that correspond to the external model is found. Although all three changes cannot be considered as significant, as a result of which the hypotheses IIa-1, IIa-2, and IIb-1 will have to be rejected, the outcomes do indicate that some qualitative rather than quantitative improvement is being made, assuming that the external model can be considered as more valid, appropriate, and/or useful than the individual participants' conceptualization on the pretest. Why participants may have decreased their domain-specific entries on the posttest (as compared to the pretest) will be discussed in 7.3.6 to start with and in more detail in Chapter 8, where the outcomes of the present study will be interpreted.

### *Assessing the effects on the basis of the questionnaire*

To assess as to whether participants believe the participative policy modelling method enabled them to change the domain-specific content of their conceptualization, they were asked to answer questions number 28, 30, 31, and 32 of the questionnaire (cf. Appendix 3). The answers to these questions have been summarized in figure 7.5 (a five point scale was used ranging from 1 to 5):

Figure 7.5: Domain-specific knowledge (questionnaire-based)

| | intention to treat | on-treatment |
|---|---|---|
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) |
| general | 2.67 (.50) | 2.56 (.42) |
| endo-exogeneity | 2.59 (.73) | 2.67 (.77) |
| multi-disciplinarity | 2.46 (.60) | 2.44 (.62) |
| N | 22 | 18 |

As shown in figure 7.5, on each of the 3 items, average scores of about 2.5 were found, indicating that the participants felt that only a very limited increase in domain-specific knowledge had been brought about by the participative policy modelling method. This conclusion holds for both the intention to treat and on-treatment group. Note that the variable 'multi-disciplinarity' has not been removed from the analyses based on the questionnaire (in contrast to the analyses based on the cognitive maps). As illustrated by its relatively low score, participants did not feel the program had succeeded in bringing about an increase in the 'multi-disciplinary' character of their knowledge.

From a general point of view, the outcomes of both evaluations (participant-based and cognitive map-based) thus seem to be in line with each other. Both methods of evaluation show that hardly any change in the domain-specific content of the participants' conceptualization of the problem has been brought about by the participative policy modelling method.

However, as already shown in the general description of the cognitive maps, the kind of concepts and relationships that are used on the posttest vary dramatically from the kind of concepts and relationships that are used on the pretest. However, since this change is not reflected in for instance the number of endogenous and exogenous concepts, those scores are not affected by it - the change cannot be made visible using the domain-specific knowledge indicators.

Having discussed the outcomes on the basis of which it was concluded that no significant individual change in domain-specific knowledge was found, account can be taken of the results of the present study with respect to it's third research question; the question concerning the *inter-individual* change in *strategic* knowledge.

### 7.3.3 INTER-INDIVIDUAL STRATEGIC KNOWLEDGE (III)

The third research question concerns the effects that are brought about by the participative policy modelling method on the degree to which participants are alike in the strategic knowledge they use when conceptualizing a particular policy problem. As such, it aims to evaluate the degree to which a shared understanding with respect to the way in which the participants organize their domain-specific knowledge when considering a complex, ill-structured problem such as the cost development of the Dutch Health Care system, has been brought about.

In order to arrive at a measure of the degree to which the participative policy modelling method has been able to bring about such a shared understanding, use will be made of the same variables that were used to assess the individual change in strategic knowledge: the number of feedbackloops, the connectivity of the conceptualization, its chaining, and the number of time-indications used by the participants. However, rather than focusing on the differences in means between the pre- and posttest scores, an assessment will be made of the differences in variance between the pre- and posttest. To assess as to whether a substantial difference in variance (standard deviation) can be found between the pre- and posttest, use has been made of the Bartlett test in relationship to the following hypotheses:

*Hypothesis III-1: on average, the variance of the* ***feedbackloop*** *variable will be smaller on the posttest than on the pretest.*

*Hypothesis III-2: on average, the variance of the* ***connectivity*** *variable will be smaller on the posttest than on the pretest.*

*Hypothesis III-3: on average, the variance of the* ***chaining*** *variable will be smaller on the posttest than on the pretest.*

*Hypothesis III-4: on average, the variance of the* ***time-indications*** *variable will be smaller on the posttest than on the pretest.*

Application of the Bartlett test resulted in the following outcomes:

Figure 7.6: Inter-individual change in strategic knowledge

| Strategic | intention to treat | | | on-treatment | | |
|---|---|---|---|---|---|---|
| | pretest | posttest | | pretest | posttest | |
| | sd | sd | p-value | sd | sd | p-value |
| feedbackloops | .67 | .86 | .27 | .71 | .86 | .44 |
| chaining | 1.73 | 1.99 | .52 | 1.86 | 2.07 | .68 |
| connectivity | .15 | .12 | .31 | .15 | .11 | .31 |
| time-indications | 1.04 | .57 | .01 * | .92 | .59 | .08 * |
| N | 22 | 22 | | 18 | 18 | |

In figure 7.6, an overview is given of the variance that was found for each of the strategic knowledge variables on both the pre- and posttest. With respect to the variable 'feedbackloops', for instance, the variance that was found on the pretest is equal to .67, whereas on the posttest a variance of .86 was found for the intention to treat group.

Since the participative policy modelling method is expected to bring about a reduction of the variance , the change in variance of the variables 'number of loops' and 'chaining' clearly is opposite to what was expected (this goes for both the intention to treat and on-treatment group). The variables 'connectivity' and 'time-phase relationships', by contrast, do show a change in the right direction. The decrease in variance of the time-phase relationship can be considered as significant (p=.01 for the intention to treat group, and p=.08 for the on-treatment group). However, because the decrease in *variance* of the number of time-indications variable is joined by a substantial decrease in the average *number* of time-indications, the shared understanding arrived at seems to be the result of a 'collective' agreement among the participants not to use time-indications any more. As explained before, an increase in shared understanding can only be concluded to if the variance (standard deviation) of a particular variable is reduced, *and* the average score on the very same variable has not been decreased (ideally, it has been increased rather than remained constant). Since the 'time-phase relationship' variable is unable to meet these two requirements, the reduction of variance cannot be considered as an increase in shared understanding. Hence, all four hypotheses will have to be rejected, including Hypothesis III-4, regarding the variance of the variable 'number of time-indications'.

Since no item has been included in the questionnaire to assess the degree to which participants were themselves of the opinion that a change in their shared strategic understanding had been brought about by the participative policy modelling method, no comparisons between the questionnaire outcomes and the results based on the cognitive mapping approach can be made.

### 7.3.4 INTER-INDIVIDUAL DOMAIN-SPECIFIC KNOWLEDGE (IVa,IVb)

#### *Assessing the effects on the basis of cognitive maps*

Having discussed the outcomes related to inter-individual changes in the strategic dimension of one's conceptualization, account should be taken of the degree to which the participative policy modelling has succeeded in bringing about a kind of shared understanding, or increase in the homogeneity or commonality among the participants with respect to the domain-specific knowledge they use to refer to the problem at hand. To examine the commonality among the participants, attention will be paid to the variances of the variables that fall under the domain-specific knowledge dimension, for it is assumed that a decrease in variance (standard deviation) can be considered as an increase in commonality. To assess as to whether a significant difference in variance can be found, use will be made of the Bartlett test, testing for differences in dispersion. The hypotheses that will be tested with respect to the participants domain-specific commonality are:

*Hypothesis IVa-1: on average, the variance of the* ***endogeneity variable*** *will be smaller on the posttest than on the pretest.*

*Hypothesis IVa-2: on average, the variance of the* ***exogeneity variable*** *will be smaller on the posttest than on the pretest.*

*Hypothesis IVb-1: on average, the variance of the* ***external model*** *variable will be smaller on the posttest than on the pretest.*

The outcomes of the Bartlett tests carried out to test these three hypotheses, can be depicted as follows:

| Figure 7.7: Inter-individual change in domain-specific knowledge | | | | | | |
|---|---|---|---|---|---|---|
| Domain-specific | intention to treat | | | on-treatment | | |
| | pretest | posttest | | pretest | posttest | |
| | sd | sd | p-value | sd | sd | p-value |
| endogeneity | 18.66 | 12.83 | .09 * | 19.43 | 13.62 | .15 |
| exogeneity | 6.55 | 9.11 | .14 | 7.19 | 9.24 | .31 |
| external model | 7.48 | 9.16 | .36 | 8.02 | 9.81 | .41 |
| N | 22 | 22 | | 18 | 18 | |

Figure 7.7 shows that with the exception of the standard deviation of the variable 'endogeneity' at the intention to treat group (p=.09), none of the variances of the pre- and posttest measures differ significantly. Hence, it cannot be concluded that the participants have become more alike in their conceptualization of the problem due to the participative policy modelling method: hypotheses IVa-1, IVa-2, and IVb-1 all have to be rejected.

Moreover, only with respect to the endogeneity variable a decrease in variance is found. However, since the average score is decreased, it fails to qualify as an increase in shared understanding. The other two variables (exogeneity and external model) show an increase rather than a decrease in variance. As a consequence, no shared understanding has been brought about with respect to those variables as well. It is important to note that there may well be a relationship between the increase in the average scores on these two variables and the increase in variance found at the very same variables. The same relationship probably holds for the decrease of the endogeneity score and the decrease in variance.

### *Assessing the effects on the basis of the questionnaire*

The only item that was included in the questionnaire to assess how the participants felt about the inter-individual change in domain-specific knowledge brought about by the method, was question number 35 (cf. Appendix 3). With respect to this particular question, an average score of 2.34 (standard deviation:1.03) was found for the intention to treat group, whereas the on-treatment group on average scored 2.50 (standard deviation: 1.01). Both values indicate that some increase of the shared understanding among the participants was arrived at, although this increase is only a moderate one.

As such the outcomes of the cognitive mapping approach are in line with the findings based on the questionnaire: both suggest that no significant improvement in shared understanding has been arrived at by the participative policy modelling method. As a consequence, the hypotheses outlined at the beginning of this section will have to be rejected.

Having answered the research questions that focus on the degree to which the participants change the knowledge they have (or use) with respect to a particular policy *problem* at hand, account will be given to the fifth research question in which knowledge of the other *participants* will be focused upon, that is, knowledge of how those participants look upon the very same problem. This 'awareness of the others' point of view' will be dealt with in the next section.

## 7.3.5 AWARENESS OF THE OTHERS' POINT OF VIEW (V)

To assess as to whether participants have become aware of each others' point of view, participants were asked to answer questions number 17, 24, 25, 33, 34, 42, and 43 (cf.

figure 4.7a). Since the evaluation of the effects of the participative policy modelling program on the awareness they have of the other participants, will be based on the questionnaire only (no indicators based on the participants' cognitive maps have been included in the present study), the hypothesis that will be formulated with respect to the fifth research question, will not be tested formally, that is, on the basis of differences between pre- and posttest.

*Hypothesis V: on average, participants will have increased their **awareness of the other participants' point of view** as a result of taking part in the participative policy modelling method*

The descriptive statistics for the items that make up the 'awareness of others' point of view', are presented in figure 7.8:

| Figure 7.8: Awareness of others' point of view | | | | |
|---|---|---|---|---|
| | intention to treat group | | on-treatment group | |
| | $\bar{x}$ | sd | $\bar{x}$ | sd |
| v17 | 3.61 | .64 | 3.70 | .62 |
| v24 | 3.48 | .59 | 3.75 | .60 |
| v25 | 3.71 | .96 | 3.70 | .80 |
| v33 | 2.12 | .73 | 2.14 | .66 |
| v34 | 2.48 | .87 | 2.43 | .87 |
| v42 | 1.91 | .83 | 2.00 | .84 |
| v43 | 3.19 | 1.08 | 3.22 | 1.11 |
| awareness | 3.24 | .32 | 3.28 | .33 |

It is shown in figure 7.8 that on average a value of 3.28 for the on-treatment group and a value of 3.24 for the intention to treat group is found, indicating that the participants are of the opinion that some increase in the awareness they have of the other participants' point of view has been arrived at. Hence on a general level, it seems as if the participative policy modelling method has been able to meet it's fifth objective: bringing about an awareness of how other participants look upon the problem.

On a more detailed level, however, some striking differences are found in the answers given to the items that make up the 'awareness of others' point of view' dimension. Particularly striking is the fact that when the participants were asked to state as to whether the participative policy modelling method enabled them to increase their knowledge of how their colleagues think about the health care system, the scores that relate to individual sessions (questions number 17, 24, 25, and 43[126], with the exception of question 42), are found to be substantially higher than the scores on the items that assess the effects of the program on all three session taken together (questions number 33 and 34). A possible explanation for this difference is that having stated that no real change had been brought about by the first three sessions with respect to the domain-specific and

[126] Note that on question 43, asked at the end of the fourth session, a somewhat lower score is found than the average scores on question number 17, 24, and 25. The reason for this probably is that, although quite some interaction took place during the final session, no exciting new points were raised by any of the participants, for the fourth session mainly consisted of a recap of the issues that had been dealt with in the first three sessions and a brief discussion of the implications of these issues for the existing organizational policy document.

strategic knowledge domains (e.g. by answering questions number 26 to 32 of the questionnaire), participants may have gotten in the habit of rating the items somewhat low and may have answered the items that followed (i.e. questions number 33 and 34) accordingly. One reason why it is expected that an instrumental contamination like this may have taken place, is because much higher scores were found for the items that assessed the degree to which an increase in awareness had been brought about by the *individual* components of the *individual* sessions.

### 7.3.6 SUMMARIZING THE EFFECTS

The outcomes of the present study with respect to the five basic research question can be summarized as follows:

| | improvement | |
|---|---|---|
| Question I: | individual strategic knowledge | no |
| Question IIa + IIb: | individual domain-specific knowledge | no |
| Question III: | inter-individual strategic knowledge | no |
| Question IVa + IVb: | inter-individual domain-specific knowledge | no |
| Question V: | awareness of others' point of view | somewhat |

As shown above, no changes were found to be brought about by the participative policy modelling method[127] except for the increase in awareness of the others' point of view. Moreover, the differences that were found not only proved to be not significant, but some of them were also in the wrong direction, suggesting that taking part in a participative policy modelling program would result in a decrease rather than an increase in strategic and domain-specific knowledge (this applies to both individual and inter-individual knowledge).

Since the outcomes clearly contradict the expectations outlined in Chapter Three (the expectation that participative policy modelling is able to bring about a *cognitive* change at those participating in it), account will be given of the reasons why participative policy modelling seems to be unable to live up to its expectations. Hence, prior to the presentation of how the participants felt about the program they took part in (based on the questionnaire administered at the end of each session), and the discussion of the most important specifying effects of the variables included in the questionnaire, some of the factors that may account for the somewhat surprising outcomes of the present study will be discussed. To do so, and anticipating a critical reflection of the study in Chapter 8, the issue of context-dependency or situationality (cf. Chapter 2, section 2.6, and Chapter 4 section 4.2) will be re-addressed. It is believed that changes in the context of those participating in the participative policy modelling sessions have affected the meaning and usefulness of the program offered as part of the present study substantially, and consequently have affected the influence participative policy modelling has had on the conceptualization of those taking part in it.

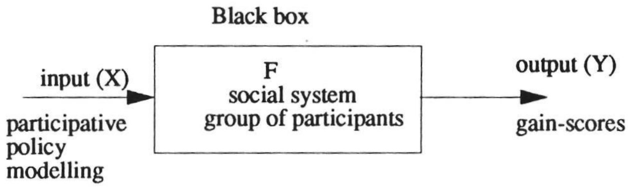
To account for the surprising results of the present study, one may have to take an insider's perspective in addition to the traditional observer's point of view (Klabbers, 1990). In order to understand what has taken place during the participative policy modelling sessions and why no major changes were found, one should not look upon the method and the people taking part in it from an outsider's point of view only but also consider the process from what can be called an insider's perspective. In addition to this, one may have to change one's perspective every now and then, to increase one's understanding of the effects brought about by the input (participative policy modelling) and the processes that can account for these effects.

---

[127] Because no significant results were found with respect to both domain-specific and strategic knowledge, no interesting observations could be made about the relationship between the two.

If we look at the evaluation of the participative policy modelling methods from what can be called an observer's point of view (Klabbers, 1986; 1988; 1990), participative policy modelling seems to be unable to bring about a change in the way in which those who took part in the programme look upon the policy problem focused upon in the sessions. This because no significant differences are found between the post- and pretest results - no significant improvement is made in terms of strategic and domain-specific knowledge. The observer's point of view can be represented as follows:

Figure 7.9: Black box




According to the black box model[128], it is on the basis of a perceived correspondence between in- and output that knowledge of the transformation function F can be arrived at. Moreover, provided the observations have been carried out scientifically, it is assumed that the knowledge thus arrived at is context-independent, predictable, and applicable to different situations and points in time. As such, the model reflects a kind of rationalistic conception of reality in which context-independent rules and regularities are being looked for.

If we look at the outcomes of the present study from an observer's point of view, we clearly would conclude that participative policy modelling program is not capable of bringing about a significant change in the way in which participants look upon the problem, because the posttest scores do not differ significantly from the pretest scores. The system (or black box) that is responsible for transforming input into output (i.e. the group of policy makers taking part in the program) seems to act in such a way that no change in conceptualization is brought about as a response to taking part in the participative policy modelling sessions.

However, as mentioned above, to account for the effects brought about by the participative policy modelling method, a closer look at the system (transformation function) responsible for the absence of any effects may have to be taken. One should try to open up the black box and look at the processes that take place in it. In addition to focussing on the processes that are taking place, that is, the behaviour of the system over time, one should change one's perspective from an outsider's perspective to an insider's perspective every now and then.

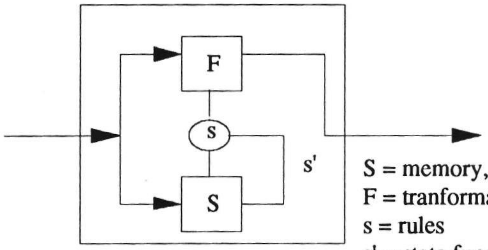
The reason why such an insider's perspective seems to be promising to account for the (lack) of effects, is that, as with all systems in which human beings play a part, one cannot understand the dynamic behaviour of those systems (e.g. the effects brought about by it) on the basis of observed sequences of input/output pairs only. This because the internal state of the system (the people that take part in the program, their values, expectations, knowledge etc.) affect the system's output as well. Stated differently, different groups participating in the very same participative policy modelling program can and probably

[128]The black box model is also referred to as the Trivial Machine, a black box F processing input X into output Y (Von Foerster, 1984; Klabbers, 1988; 1990). As soon as the X-Y correspondence through F is established, any X will generate a specific Y, independent of time. Hence Trivial Machines are predictable, history independent, synthetically deterministic, and analytically determinable (Klabbers, 1990). It follows the rationalistic conception of reality according to which knowledge is composed of abstract, context-independent, formally interconnected domain-specific concepts.

will bring about different effects - there is not a single transformation function F that will univocally determine what the group of participants' reaction to the program will be.

Hence, to fully understand the relationship between participative policy modelling and the effects produced by it, an insider's point of view seems to be needed as well to take into account the historical or situational aspects of the group of people taking part in the program, to consider these aspects of the process of transforming a particular input into a particular output. The model that aims to take these processes into account is called the Non-Trivial Machine (NTM) and can be represented as follows:

Figure 7.10: The Non-Trivial Machine




S = memory, history
F = tranformation function
s = rules
s' = state function

*Source: Klabbers, 1991, p. 212*

The model of a Non-Trivial Machine is useful in that it helps us to look at the effect of participative policy modelling from an insider's point of view, making is visible that the system's (NTM) internal state is affecting the processing of inputs into outputs (F) which in turn affect the internal state as well. Groups of people can be considered as Non-Trivial Machines, for the outcomes of their behaviour, the processing that is taking place, is not the result of an invariant F, but is also affected by the (historical) status of the group. Depending on the group's background, interests, values, and knowledge, different effects on the processing functions (F) will be found.

This brings us to two additional but related reasons why the historical situation and the changes in the system's internal state seem to be important to take into account to understand and explain the outcomes acquired with an observer's point of view. Ther first reason concerns the context within which participative policy modelling is applied in the present study. The second the nature of the output variable(s) this study is dealing with.

As far as the context within which participants are exposed to participative policy modelling is concerned, recall that it is the present study's aim to assess the effectiveness of participative policy modelling as a tool to support policy makers in the process of policy-making when dealing with ill-structured problems. In Chapter Two, it was stated that ill-structured policy problems are both cognitively and socially complex. Because of this complexity, conflicting values, competing claims, and different perspectives on the problem arise - the policy-making process is a process moulded by personal and organizational characteristics, loyalties and interests. Consequently, policy-making support is taking place in a *context* which is not purely knowledge-driven, but also affected by personal and political factors. By taking an insider's perspective, some of the processes that have taken place and the meaning of these processes which have affected the (cognitive) outcomes of participative policy modelling can be made visible. As Morecroft (1992, p. 25) puts it: *"It is easy enough for modellers to become fascinated with representation scheme's, software, the mental models of clients, and cognitive imperfections in dynamic decision making, yet lose sight of individual and group behaviour on which models depend for both ideas and legitimacy"*. Knowledge production is not taking place in a vacuum, independent of inter-personal processes and external developments, and can only be understood in relation to those factors.

With respect to the nature of the outcome variable(s), recall that it has been the objective of the present study to assess the participative policy modelling's potential to contribute to the policy-making process in general, and due to the nature of the problems that are focused upon, its potential to contribute to the conceptualization of policy problems in particular. In other words, the present study has been concerned with the evaluation of the method's potential as an aid for improving the conceptual orientation of social systems or as Klabbers (1986) puts it, *to raise the general level of comprehension of these systems* [i.e. groups of people]. Because the effects of participative policy modelling have been defined in terms of cognitive changes[129], an insider's perspective should be taken as well, for "*knowledge is situated, being in part a product of the activity, the context, and culture in which it is developed and used*" (Brown, Collins, and Duguid, 1989, Klabbers, 1990). Changes in the social system's cultures and values (it has already been stated that the technological stratum of the health care system has remained relatively stable throughout the whole program, no major changes in the material structure of the system (patient flows) were found), however, do affect the usefulness of the knowledge provided in the preliminary model, the participants' conception of the program, their willingness to take part, the degree to which they aimed to re-structure the program themselves, and the meaning that was given to all this. Hence, to fully understand the outcomes of the present study, account should be given to the processes that were taking place during the sessions, the context or historical situation that affected the participants behaviour and the consequent meaning participants have attributed to the program.

Summarizing the above, in order to understand the effects participative policy modelling has had on the group of people taking part in it, one should not only focus on differences between pre- and posttest scores, for we are dealing with a social system that is capable of reacting different to one and the same input depending on its initial values. The people's history and the context in which the sessions are taking place have a large impact on their outcomes. Moreover, since we aim to support the policy-making process, issues such as power, culture, values and norms should be taken into account when explaining the outcomes of the process, that is, when trying to find the transformation function F to account for the effects of the participative policy modelling method. This because the meaning, usefulness, and construction of knowledge is heavily affected by its context.

By focusing on the patients flows, we aimed at providing the participants with a preliminary model that would not be too sensitive to changes in the context of the sessions (i.e. become outdated). As discussed in Chapter Four, the causal or technological stratum, represented in the preliminary model proved to be stable in that almost all of its causal relationships still hold today, three years after the start of the construction of the preliminary model, and two years after its completion.

However, as far as the structural (representing the formal and informal organization), and cultural (the norms and values)[130] aspects of the social system taking part in the sessions are concerned, some important changes in context took place at the time the sessions were held. Changes that are likely to have affected the transformation function F and consequently the outcomes of the participative policy modelling sessions[131].

---

[129]In the next chapter the present study's restriction to cognitive effects will be reviewed critically, in relationship to the methodological implications of adopting a Non-Trivial Machine perspective when examining the effects of a policy making method such as participative policy modelling.

[130]For a more detailed description of the three aspects or layers of social systems, see Klabbers (1985).

[131]Traditionally, the influence of a particular context on the outcomes of an experiment such as the one carried out in the present study, is referred to as 'history' (Campbell & Stanley, 1966; Cook & Campbell, 1979). According to this notion of context, events that take place in between the pre- and posttest may affect the outcomes of the experiment.

At first sight, the lack of effects found in the present study thus seem nothing more than an indication of the possibility that the outcomes of a study are threatened by changes that have taken place in the context

One of the important aspects of the context in which the participative policy modelling sessions were carried out concerns the *reasons why the organization was willing to participate.* As described in Chapter Five, the health care insurance organization had felt the need to prepare itself for a more active policy making role in the near future. Moreover, it wanted to use the sessions to increase the commonality among the people from the Medical Department and the Financial-Legal Department, that is, to arrive at a kind of shared understanding of the Dutch health care system, shared by both departments. However, some disagreement among the participants existed as to whether participative policy modelling was the most appropriate way of establishing these objectives. Particularly with respect to the second objective, the integration of the Medical and Financial-Legal Departments, some people from the Financial-Legal Departments expressed the feeling that since they had already started working together more closely with the people from the Medical Department, there no longer was a need to take part in a program by means of which the communication and cooperation between the two departments could be improved. One member of the Medical Department withdrew from the program after the first session because he felt there no longer was a need to take part in it.

Much of the difficulty we had in getting all members of the two departments participate enthusiastically in the program was because of the fact that the participating organization had two rather than one single objective for taking part in the program. Part of the people participating in the program looked upon the participative policy modelling method as a means to increase the organization's policy making support function. Others, however, expected the program to support and improve the communication between the two departments. Although these objectives need not be in conflict with each other, they turned out to be somewhat difficult to combine. After the first session for instance, some of the people complained about the length of the first preparatory text for they felt that reading a text for about four hours on an individual basis was not doing any good to their communication[132]. The people who expected to learn about the health care system to prepare themselves for a more active policy making role however, felt the workbook had been very useful and the session could have focused upon some aspects of the model in more depth.

---

of the experiment The solution to this would be to isolate the experiment from its context or to make sure that both groups (in case a control group design is used) experience the very same context

On second thought however, the relationship between context and conceptualization or knowledge of a (policy) problem differs from the threat to internal validity called 'history' in that the context when dealing with participative policy modelling cannot and should not be excluded from the processes that take place during the sessions and the effects that are measured on the posttest History should not be considered as a negative factor disturbing the outcomes of the study, it is the factor without which no meaningful changes in the participants' conceptualization can be brought about, for it is only within a particular context (people, situation, and time) that the knowledge materialized in the conceptual and computer model does make sense Changes in context hence deeply affect the meaningfulness and usefulness of this knowledge, but cannot and should not be excluded from the design In the next chapter, the methodological implications of this context-dependency will be elaborated upon on more detail, for it not only affects the knowledge produced during the participative policy modelling sessions, but the knowledge that can be obtained from the present study itself as well

[132]At one stage, the night before the second session was about to take place, the CEO of the Division 'Care' [Zorg], called to explain that some of the participants wanted to step out of the program because they felt the program was not capable of living up to their (communication support) expectations, and because the after-effects of the merger were taking too much extra time to justify spending time on non-work related activities such as the participative policy modelling program The CEO subsequently expressed that he had talked those people out of the idea of withdrawing from the program and that he would personally see to it that all participants would complete the program The reason why he informed us about this situation was because he wanted us to know that the organization was having some difficulties in coping with the after-effects of the merger as a result of which some of the people participating to the program would not be taking part as enthusiastically as they would have done under normal situations (it turned out that two members of the Board had announced their withdrawal at the day the CEO gave us a call)

To account for this combination of objectives, recall that the organization had been involved in a *merger process*, as a result of which the originally felt need for policy-making support had changed partially into a need for communication support. Many of the people who felt the program served to improve their communication, looked upon the health care model as an interesting way to start communicating with each other rather than as a way to improve and unify their conceptualizations of the health care system. As a result, one should not be too surprised not to find any major changes in the conceptualization of those taking part in the program. Changing from a policy-making support to a communication support point of view clearly affects the way in which people look upon the applicability and usefulness of the knowledge provided and created during the sessions.

Besides the merger process the organization had been involved in (unifying two organizations and two departments), the outcomes of the participative policy modelling were also affected by the *work-pressure* that was felt by the participants from the Financial-Legal Department due to the rapid changes in the financing structure of the health care system. One of these changes for instance, concerned the introduction of a budgeting system for regional health care insurance organizations. In order to arrive at such a budgeting system (and an idea of what these budgets would look like), regional health care insurance organizations were asked to come up with figures concerning their next year's budget (and spending) at the end of 1990. Without any experience in doing so, regional health care insurance organizations were asked to come up with figures that had to be a relatively accurate estimate of the health care consumption in their region for the year to come. Since the budgets had to be ready before January 1st 1991, it should not come as a surprise that many of the participants from the Financial-Legal department (responsible for the estimates) felt activities other than the participative policy modelling program were important as well.

In sum, it has been argued that the changes that have taken place in the context of the participative policy modelling program may account for the lack of outcomes reported above. The changes in the context in which the present study has been carried out may have affected the perceived usefulness and applicability of the program, as a result of which the transformation function (F) may have changed. In other words, by looking more closely at the phenomenon under consideration, the transformation function F appears to be more complicated than the invariant one that is being conceived from an observer's point of view, illustrating the context-sensitivity of supporting the policy makers' conceptualization of a particular policy making problem. One cannot simply conclude that participative policy modelling is not capable of bringing about any changes in the conceptualization of its participants by taking an observer's point of view, for the transformation function (how people respond to the program) clearly is affected by the situation in which the program is carried out and the people taking part in it. Hence, in assessing the effects of participative policy modelling, it seems promising to look upon the processes that have taken place and the products these processes have resulted in from both an outsider's and an insider's point of view. For this the metaphor of the Non-Trivial Machine can be used, for it assists in focusing on the importance of the social system's initial values in the process of transforming input into output[133] and

[133]In studying social systems, the Non-Trivial Machine approach should be generalised by replacing the state function S by actors and rules, who represent the state function of the system. By doing so, one switches from a machine to an actor approach: "*In the actor system, via communication and coordination, actors develop rules and procedures both for communicating and for intervening in their internal environment F. Through this recursive processing of information i.e. conversation and communication they continuously transform their collective structure. Actors engage in a process of social construction of reality and their collective awareness of this process is an indicator of the internal state of the system*" (Klabbers, 1990, p.5). The notion of an actor system was introduced because it is believed that the initial values of social systems and the social system's transformation rules are hardly ever known (Klabbers, 1988) and that they can be changed, for instance in the process of policy making (Klabbers, 1986). Depending on the situation, that is depending on the way in which the actors interact, the way in which they relate to a

prevents us from trivializing the behaviour of the system by reducing it into observed in- and output pairs.

Thus by changing from an observer's point of view to an insider's point of view, additional understanding of the reasons why the participative policy modelling program has not been able to bring about any significant changes in conceptualization, can be acquired. This understanding will prevent us from prematurely concluding that the transformation function F was found to be such that 'under all situations', participative policy modelling would be unable to live up to its expectations. Moreover, the change in perspective has made us realize that, from both a theoretical (i.e. concerning the 'what' of research) and methodological (i.e. concerning the 'how' of research (Bredemeyer & Greenblatt, 1981)) point of view, the effects of taking part in participative policy modelling may have to be defined differently so that account is taken of its inherent context-sensitivity. In other words, the change in perspective has made us realize that a pretest-posttest design may not be sufficient to evaluate the effects of participative policy modelling. However, it may also be that the benefits brought about by the participative policy modelling method will be different than the ones discerned in the present study. In the next chapter, the theoretical and methodological implications of such a change in perspective for the evaluation of the effects of participative policy modelling will be discussed in more detail and some recommendation for evaluations to come will be made.

## 7.4 POTENTIALLY SPECIFYING VARIABLES

Having presented the outcomes of the present study with respect to the five basic research questions and having given some explanation of the reason why they are what they are, an overview will be given of the way in which the participants looked upon the program and the amount of time that was spent on it. Since the background variables have already been discussed in Chapter Five to provide an idea of what the people who participated to the program look like, the presentation of the potentially specifying variables will be restricted to the participants-based evaluation of the program (in terms of both components and aspects of the program) and the amount of time that was invested in it by the participants.

### 7.4.1 EVALUATION OF THE PROGRAM: COMPONENTS AND ASPECTS

As explained in more detail before, the participants were asked to express how they felt about the program by means of a questionnaire that was given to them at the end of each session. The outcomes of this self-evaluation of the program in terms of both components and aspects can be presented as follows:

particular reference system (e.g. the health care system), and the way in which they interpret the system's rules, actors transform the social system they are a part of, thereby affecting the output the system is capable of bringing about (e.g. the system's reaction to participative policy modelling). In Chapter 8, the notion of an actor system will be elaborated upon.

| Figure 7.11: Participant-based evaluation of the program on treatment group | | | | | |
|---|---|---|---|---|---|
| | session 1 | session 2 | session 3 | session 4 | total score |
| | $\bar{x}$ (sd) | $\bar{x}$ (sd) | $\bar{x}$ (sd) | $\bar{x}$ (sd) | $\bar{x}$ (sd) |
| *components* | | | | | |
| workbook | 3.22 (.35) | 2.98 (.42) | 3.47 (.52) | not included | 3.18 (.35) |
| small-group activities | 3.91 (.32) | 3.72 (.42) | 3.55 (.58) | not included | 3.77 (.29) |
| plenary session | 3.41 (.57) | 3.44 (.54) | 3.56 (.53) | not included | 3.48 (.36) |
| *aspects* | | | | | |
| useful | 3.68 (.59) | 3.48 (.50 ) | 3.65 (.60) | 2.44 ( .96) | 3.54 (.38) |
| learned from it | 3.28 (.53) | 3.16 (.55) | 3.43 (.47) | 2.88 (1.03) | 3.10 (.37) |
| interesting | 4.16 (.43) | 3.79 (.38) | 3.83 (.42) | 2.50 (1.10) | 3.80 (.47) |
| easy | 2.82 (.61) | 2.91 (.46) | 3.25 (.38) | 3.19 ( .91) | 3.04 (.12) |
| length | 3.17 (.58) | 3.15 (.74) | 2.82 (.34) | not included | 3.13 (.54) |

The scores are based on the on-treatment group for it is felt that only those who have taken part to the program, are in the position to evaluate the program[134].

### *Components*

As far as the average scores of the components 'workbook', 'small-group activities' and 'plenary session' are concerned, it is clear that the small-group activities have been appreciated most. Their average score is substantially higher than the scores given to the other two components. Note that the preparatory texts (workbooks) were enjoyed least, their score is only just above 3 (on a scale from 1 to 5). As a matter of fact, some of the participants did complain about the amount of time that was required to prepare reading the workbooks for not only were they hard pressed for time, they also felt that 'individual' participation in the model-building project was not the most appropriate way to arrive at a shared understanding. Except for the somewhat low scores on the preparatory texts (workbooks), one can legitimately say that the other two components (i.e. the small-group activities and the plenary sessions) were evaluated positively.

Regarding the evaluation of the components over time, that is, the values that were given to each of the components on a particular session, only values concerning the first three sessions can be found. The reason for this is that no workbook was given to the participants to prepare for the fourth session, and that no small-group activities were carried out by the participants in the fourth session. The activities that took place in the fourth session, were carried out in a plenary mode only.

On average, no major differences are found in the way in which the participants looked upon the plenary session - the ratings given to all three plenary sessions did not vary much. A somewhat different pattern was found for the workbooks given to the participants prior to the sessions. The second workbook, discussing potential exogenous influences, was not received too well, compared to the first and third workbook. As far as the small-group activities are concerned, note that the first small-group session was

---

[134]No significant differences were found between the people who qualified for the intention to treat group but failed to qualify for the on-treatment group with respect to background variables and evaluation of the program. The only variable where a significant difference was found is the amount of time spent on the program. Since the on-treatment group was selected on the basis of this variable, a significant difference with respect to the amount of time invested in the program was to be expected.

received best (in this session, the structure and dynamics of the health care system was discussed), whereas the third session was rated worst (still substantially higher than 3.0 though, indicating that on the whole people did appreciate the small-group activities). Two reasons why the second and third sessions were rated lower than the first come to mind. Firstly, it is possible that the first was rated higher because it was the first of four sessions, as a result of which both content and approach (working in small groups, having plenary presentations) were new to the participants. However, the same pattern is not found for the scores on the other two components, workbook, and plenary session. Secondly, and probably more important is that the small-group activities in the second and third session had to be rushed a little bit because of lack of time. Participants had to be asked to stop their discussions or computer runs and start the plenary presentation of their observations. This may have affected the ratings of those sessions somewhat. Since no attempt was made to evaluate the participative policy modelling program carried out in terms of its components in the Vennix (1990) study, no comparisons between the present study and the Vennix (1990) study can be made with respect to participant-based evaluation of the workbooks, plenary sessions, and small-group activities. Besides, because the two studies differ substantially in the program offered to the participants, the usefulness of comparing their components can be questioned.

### *Aspects*

The average aspects scores vary from 3.0 (for the aspect 'easy') to 3.8 (for the aspect 'interesting'). Because people consider the program to be interesting and useful on average, despite the fact that they claim not to have learned much from it, the question arises what it is that has made the program useful and of interest to them. The correlations between the aspects 'useful' and 'interesting' on the one hand, and the aspect 'learned from it' on the other hand[135], do suggest that the people who claim to have learned from it, also consider the program to be useful and interesting. However, because of the differences in averages, the possibility exists that the participants either apply a different scale for the aspects 'useful' and 'interesting' than they do for the aspect 'learned from it', or have additional reasons to look upon the program as interesting and useful. What these additional reasons may be, will be discussed in the next chapter, where the effects brought about by the participative policy modelling program will be interpreted and discussed in more detail.

Moreover, notice that on average, people have found the participative policy modelling sessions somewhat difficult, particularly the first two sessions concerning the structure and dynamics of the model (including the influence exogenous factors are having on the behaviour of the model over time). The third session however, in which the participants were asked to examine the effects of potential solutions (assess the effects of potential policy measures) was considered to be somewhat easier than the previous two sessions. The relatively low score on the 'easiness' aspect corresponds to the difficulties many of the participants were having in answering the questions asked about the structure and dynamics of the health care system in the workbooks that were given to them prior to the sessions. Despite the fact that most participants were able to talk about the health care system and its problems in a general way, it turned out to be relatively difficult to them to indicate which variables were connected to each other and in what way. To illustrate, in the second session, participants were asked to describe the relationship between ageing population and the average price of medicine used by the people visiting a medical specialist. Not only did the participants disagree as to whether an increase, decrease or no change at all in price was to be expected, they found it almost impossible to tell how much this change would be (e.g. an X per cent increase of the average age of the people

[135] A positive correlation is found between both 'useful' and 'learned from it' and 'interesting' and 'learned from it' ($r=.49$, $p<.05$, and $r=.47$, $p<.05$).

attending a general practitioner would result in an Y per cent increase of the average price of medicine of used by the people visiting the general practitioner).

It is interesting to see that the change in context preceding the second session that was discussed above (the withdrawal of two members of the Board), coincides with a low in the evaluation of the aspects 'useful', 'learned from it', and, 'interesting'. This suggests that indeed a relationship may exists between the way in which the program is being perceived and the changes that take place in the context of the sessions.

To compare these outcomes to the scores found by Vennix (1990, p. 173 and p. 179), use will be made of the 'useful'[136] and 'easy' aspects only. This because those are the ones that are included in both studies. The results of both studies with respect to the 'useful' and 'easy' aspects can be depicted as follows:

Figure 7.12: Comparing aspect scores: Vennix (1990, p. 173, 179) vs present study

| | Vennix | | Verburgh |
|---|---|---|---|
| aspects | students | experts | experts |
| useful | 3.66 | 4.25 | 3.54 |
| easy | 2.59 | 3.13 | 3.32 |

The results depicted in figure 7.12 show that the experts taking part to the present study correspond to the students participating in the Vennix study with respect to their opinion of the usefulness of the program. However, as far as the easiness of the program is concerned, the experts of the Vennix (1990) study and the experts participating in the present study correspond, in that they both look upon the program as not being too difficult, in contrast to the students who clearly had more difficulty in dealing with the econometric model used in the Vennix study. One possible reason for this is that students need to concentrate on both the domain-specific and strategic knowledge that is offered during the sessions, whereas it is assumed that experts do not have to pay too much attention to the domain-specific content (on which they are expected to be experts), as a result of which they can concentrate on the strategic system dynamics aspects of the program. This is not to say that the experts in the *present* study already had the domain-specific knowledge required to deal with the problem for they were dealing with an ill-structured problem. However, it is meant to say that at least they seem to have had access to a domain-specific knowledge (base) that could be used to select elements from to conceptualize and to define the problem at hand.

As far as the evaluation of the aspects over time is concerned (cf. figure 7.11), values have been arrived at for all four sessions with the exception of the aspect of 'length' (since the fourth session took only two hours, it was felt to be inappropriate to ask as to whether it had been too long).

The major difference that is found in the way in which the participants evaluate the individual sessions in terms of their aspects, is a difference that is found between the first three sessions and the fourth one. On all aspects included in the fourth questionnaire, the fourth session was rated substantially lower than the first three sessions. Participants have found the fourth session less interesting, useful, and easy than the other sessions. Moreover, they claim to have learned less from the fourth session than they had from the preceding three. The outcomes do not really surprise us for the fourth session had not been designed to improve the participants knowledge, or to be very useful and interesting. The fourth session had been developed to have the two groups of participants share their ideas, and to finish the program in a nice and relaxing way. As a consequence,

[136]Although the term 'valuable' is used by Vennix (1990) in the discussion of the results of his study (cf. Vennix, 1990, p. 173), the actual term used in the questionnaire given to his participants is 'useful' (cf. p. Vennix, 1990, Appendix 8, pp. 269-270).

it is somewhat surprising to see that the fourth session still was considered to be fairly difficult[137]. One possible explanation for this is that it was not until the fourth session that real policy-making in the sense of expressing the organization's system of values, and deciding on what to do, was taking place. Until then, the sessions had been somewhat academic in nature, being concerned with an abstract knowledge-model of the health care system in which the organization was operating. It was only in the fourth session that the participants were explicitly asked to relate the knowledge acquired in the preceding sessions to individual and organizational actions.

### 7.4.2 TIME-INVESTMENT

In the present study, participants on average spent 13.5 hours on the program. Of these 13.5 hours, 8.7 were spent on attending sessions, and 4.8 on preparation. Note that these figures concern all 26 participants. As shown in figure 7.13, these figures differ for the intention to treat and on-treatment groups because these groups were composed on the basis of the amount of time spent on the program.

Figure 7.13: Time-investment

| | N | total time invested | preparation | attendance |
|---|---|---|---|---|
| | | $\bar{x}$ sd | $\bar{x}$ sd | $\bar{x}$ sd |
| total group | 26 | 13.50 (3.47) | 4.81 (2.29) | 8.69 (2.07) |
| intention to treat | 22 | 13.76 (3.17) | 4.62 (2.06) | 9.14 (1.77) |
| on-treatment | 18 | 14.58 (2.58) | 5.14 (1.70) | 9.44 (1.69) |

If we compare those figure to the amount of time spent by the students who participated in the experimental condition in the Vennix (1990) study, we see that on preparation alone, students put in more than double the amount of time spent by the participants of the present study. On average, the participants of the Vennix study attended 93 per cent of the total 14 hours that could be spent in class, which is equal to an average of 13,0 hours (Vennix, 1990, pp. 170-171). Hence, participants to the Vennix study spent on average 22.92 hours on the program (22 hours and 55 minutes), which is substantially more than the participants in the present study did.

Another explanation of the reason why hardly any differences were found between pre- and posttest in the present study is related to the amount of time invested in the program. Getting participants to understand the dynamic behaviour of a complex system and giving them a feel for the dynamic consequences of potential policy measures is something that has proven to be very time-consuming. Klabbers (1972) for instance, had one single participant spend about six months examining the dynamics of a relatively simple system dynamics model to acquire some understanding of the model. Having taught system dynamics to groups of students, we also have the experience that getting people to understand the dynamic behaviour of a system, is a difficult thing to do, requiring a lot of effort and time on behalf of our students. Hence, one should not be too

---

[137] As a matter of fact, one of the participants stated at the plenary session held at the fourth and final session, that "*since we* (that is, the participants) *do not know for sure about the problem at hand, the only thing that we do agree upon is that we do not know much about the problem, there is hardly a point in trying to develop policies aimed at reducing the costs of health care at this stage*". "*Moreover*", he stated, "*we even do not agree upon the basic mission of our organization: is it our objective to serve our policy holders as much as possible or is it our primary objective to reduce the costs of health care for those two objectives not necessarily go together.*"

surprised not to find any major changes in conceptualization (in terms of system dynamics criteria) when participants spend on average 13.5 hours in total on the program.

## 7.5 THE SPECIFYING EFFECTS OF BACKGROUND, EVALUATION OF THE PROGRAM, AND TIME-INVESTMENT[138]

Now that an overview has been given of how the participants looked upon the program and how much time they spent on it, and using the knowledge we have about their personal background, account can be taken of the specifying effects these variables may have on the relationships between treatment and dependent variables. Recall that as far as the whole group was concerned (that is, the intention to treat and the on-treatment group), hardly any effects of the participative policy modelling method were found. The present analyses are carried out to examine as to whether differences exist between parts of the **on-treatment** group. To divide the on-treatment group into subgroups, use will be made of the potentially specifying variables. To illustrate, we will examine whether those who consider the program to be useful an those who consider it to be (relatively) useless, or those who are relatively young and those who are relatively old, show different *learning* behaviour on all five major research questions. Note that the most interesting differences only will be presented in the present section. For a more detailed description of the specifying effects of the background variables, the evaluation of the program, and the amount of time invested, the reader is referred to Appendices 12 and 13[139].

The description of the specifying effects will start with an overview of the effects brought about by the background variables and the amount of time invested in the program. Following this, the major effects of the participant-based evaluation of the program will be presented.

### 7.5.1 SPECIFYING EFFECTS OF THE BACKGROUND VARIABLES AND TIME-INVESTMENT

To assess whether any differences exist between subgroups created on the basis of the potentially specifying variables with respect to individual changes in conceptualization, differences in average gain-scores were compared using analysis of variance techniques (ANOVA and MANOVA). To assess inter-individual differences between these subgroups, differences in variance were examined using tests of Bartlett. Regarding the differences between the subgroups with respect to their awareness of the others' point of view, use was made of analysis of variance (ANOVA). The outcomes of all these analyses can be summarized in figure 7.14.

It is shown in figure 7.14, that indeed some specifying effects can be found for the present study's background variables and the variable 'time-investment' with respect to ***individual*** changes in conceptualization. The figure shows that substantial differences exist between participants who have been with the firm for a long time, and those who just recently joined the firm, participants who have a university degree and participants who have not, participants who work at the medical department and those who work at the financial-legal department, and those who spent more time on the program and those who spent less time on the program. However, with respect to the subgroups that did significantly better than their counterpart (e.g. the medical department did do better than

---

[138]In addition to these potentially specifying variables, the effects of splitting up the group in those who did best and those who did worst (not on pretest scores but on gain-scores) have been examined as well. The outcomes of these analyses are presented in Appendix 14)

[139]In Appendix 12, an overview is given of the specifying analyses carried out on the basis of the background variables and the amount of time invested in the program. In appendix 13 subsequently, a summary is given of the specifying effects brought about by the variables included in the questionnaire to evaluate the program in terms of aspects and components.

the financial-legal department), no significant improvements were found. Hence, one cannot conclude that an improvement of the domain-specific and strategic knowledge at any of the subgroups created on the basis these specifying variables was brought about by the participative policy modelling method.

Regarding the specifying effects of the background variables (and the variable 'time-investment') on the relationship between treatment (participative policy modelling) and degree of ***correspondence*** (inter-individual changes in conceptualization) among the participants, tests of Bartlett revealed that indeed some differences were found for the variables 'department' and 'education'. However in none of the four subgroups created on the basis of these two background variables, a significant reduction in variance was found. Hence, we cannot conclude that in any of the subgroups an increase in shared understanding among the participants has been brought about.

Finally, no significant differences were found between any of the subgroups created on the basis of the background variables and the variable 'time-investment' with respect to the research question ***'awareness of the others' point of view'***. As a consequence, the conclusion reached at for the whole group also holds for any of the subgroups: participative policy modelling is able to bring about a moderate increase in the participants awareness of the others' point of view.

Note that in figure 7.14, some additional information is provided for those subgroups where differences between the groups were found to exist. As part of the more detailed discussion of the effects of the potentially specifying variables in Appendices 12 and 13, account will be given of the reasons why such differences may exist, based on some of this additional information.

Figure 7.14: Background variables and time-investment **

| | *individual change* * | *perception of the modelling process* |
|---|---|---|
| Education | high education do better | higher dislike the program more; appreciate the plenary sessions more |
| Years with the firm | longer do better on domain-specific knowledge; shorter do better on strategic knowledge | longer felt the program was easier, more useful, more interesting, and more positive about plenary session |
| Department | medical department does better on domain-specific knowledge | medical department did not like the small-group activities as much |
| Gender | no differences | no differences |
| Time spent | less do better on strategic knowledge | less time spent claim to have learned from it and to have found it useful |
| Area of expertise | no differences | no differences |
| Organization | no differences | no differences |
| Policy experience | no differences | no differences |
| Age | older do better | appreciated workbooks a little bit more |

* No effects were found on inter-individual change and awareness of others' point of view
** For a more detailed description, see Appendices 12 and 13

## 7.5.2 SPECIFYING EFFECTS OF EVALUATION OF THE PROGRAM

The participant-based evaluation turned out to have only a limited specifying effect on the outcome variables of the present study. Participants who felt the program had been interesting for instance, did not have different outcome scores than the participants who felt the program had not been interesting[140]. The effects that were found, however, can be presented as follows:

| Figure 7.15: Specifying effects of the evaluation of the program** | |
|---|---|
| | *individual* * |
| program length | Those who found the program short, increased their scores whereas those who found it long, decreased their scores |
| learned from it | Those who say have learned from it, have lower strategic knowledge scores |
| useful | Those who say it has been useful, have better domain-specific knowledge scores |
| * No differences are found with respect to inter-individual changes and awareness of the others' point of view<br>** For a more detailed description, see Appendix 13 | |

As shown in figure 7.15, the variables 'length' (of the program), 'learned from it', and 'useful' were found to have some specifying effects on the degree to which ***individual changes*** in conceptualization were brought about by the treatment. The people who felt the program had been a bit long, on average decreased their scores on both the domain-specific and strategic knowledge dimensions, whereas the people who felt the program had not been too long, improved their scores on all three domain-specific knowledge related variables. Regarding the 'learned from it' variable, those who claimed to have learned from it more, surprisingly have lower scores on the strategic knowledge dimension than those who claim to a lesser extent to have learned from it. Finally, regarding the degree to which people considered the program to be useful, those who felt it had been more useful, also had higher domain-specific knowledge scores.

As far as the ***inter-individual changes*** in knowledge are concerned, a difference was found on some of the output indicators for the variables 'length', 'plenary session', 'small-group activities', and 'learned from it'. However, in none of the subgroups based on these variables an increase in shared understanding was found. Recall that to qualify as an increase in shared understanding, both a reduction in variance should have been brought about, and an increase in average gain-score should have been arrived at.

Finally, no clear specifying effects of the participant-based evaluation of the program on the ***awareness of the others' point of view*** was found. For a more detailed description of the effects that were found, the reader is referred to Appendix 13.

Now that an overview has been given of the way in which the participants looked upon the program they took part in (in the next chapter some recommendations will be given with respect to the program offered to the participants), and the most striking specifying effects of the potentially specifying variables have been presented, some comparisons with the Vennix study (1990) will be made.

---

140 A more detailed description of the effects of the potentially specifying variables on the relationship between treatment and outcome variables is given in Appendix 13.

## 7.6 COMPARISONS WITH THE VENNIX (1990) STUDY[141]

Although some comparisons with the Vennix study have been made in the preceding sections, particularly with respect to the variables that concern the participant-based evaluation of the program in terms of useful and easy, the present section will explicitly focus on the findings of the Vennix study and compare them, wherever possible, with the outcomes of the present study. Since the Vennix study concentrated on the evaluation of the *individual* effects of the participative policy modelling method, comparisons between the two studies will deal with individual changes in conceptualization only. The variables that will be included are the following[142] [143]:

V1 empirical validity (proportion of correct relationships)
= the number of correct *(model) relationships* divided by the *total number of relationships*

V2 precision of concepts (proportion of model concepts)
= the number of correct *(model) concepts* divided by the *total number of concepts*

V3 scope (proportion of concepts)
= *total number of concepts* divided by the *total number of lines*

V4 detail (relative length of paths)
= *mean length of paths* divided by the *total number of concepts* minus 1

V5 detail (relative length of loops)[144]
= *mean length of loops* divided by the *total number of concepts*

V6 integration (density)
= *total number of relationships* divided by the *total number of concepts*

V7 time factor (number of delayed relationships)
= *total number of delays* divided by the *total number of concepts*

V8 societal conditions (relative number of exogenous concepts)
= *total number of exogenous concepts* divided by the *total number of concepts*

---

[141]The reason why a separate section is devoted to this subject, is because in order to be able to compare the two studies, some adjustments to the operationalization of the present study's variables had to be made. Note that only some of the variables included in the Vennix (1990) could be arrived on the basis of the variables constructed in the present study. Despite the fact that not all variables included in the Vennix study could be constructed on the basis of the variables of the present study, comparisons between the two studies still can be made because results of individual indicators have been reported by Vennix rather than values of combinations of variables.

[142]The elements of the Vennix operationalization that are in italics, are the elements that correspond to variables used in the present study.

[143]Originally, we had in mind to compare the two studies with respect to the variable 'extension' (relative number of loops) as well. However, since only two valid cases were found for this variable in the present study, it was decided to drop 'extension' from the list of variables on the basis of which the two studies are to be compared to each other.

[144]The relative number of loops was not included because hardly any of the subjects in the present study managed to arrive at a valid score on this variables for in most cases one had to divide by zero to arrive at a score (cf. Vennix, 1990, p. 155 for a description of the way in which a score for the variable 'relative number of feedbackloops' is arrived at).

For each of these eight variables, comparisons have been made between the Vennix (1990) and the present study with respect to pre- and posttest scores and differences between these two[145]. The results of these comparisons are displayed in figure 7.16:

Figure 7.16: Comparisons with the Vennix (1990) study

| | Verburgh | | | | Vennix | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | students | | | | experts |
| | pretest | posttest | t-value | p-value | pretest | posttest | t-value | p-value | pretest |
| | x̄ sd | x̄ sd | | | x̄ sd | x̄ sd | | | x̄ sd |
| V1 | .42 (.11) | .49 (.15) | -1.62 | .06 * | .99 (.03) | .98 (.03) | -1.77 | .08 * | 1.00 (.00) |
| V2 | .58 (.10) | .65 (.12) | -1.71 | .05 * | .87 (.09) | .92 (.08) | 4.65 | .00 * | .93 (.35) * |
| V3 | 1.32 (.32) | 1.36 (.21) | - .64 | .27 | .43 (.15) | .43 (.12) | .09 | .93 | .40 (.18) |
| V4 | .32 (.13) | .32 (.09) | - .04 | .49 | .22 (.11) | .20 (.09) | 1.04 | .32 | .27 (.06) |
| V5 | .05 (.07) | .04 (.07) | .15 | .44 | .24 (.09) | .20 (.08) | -1.66 | .11 | .12 (.11) |
| V6 | 1.04 (.15) | 1.02 (.11) | .71 | .25 | 1.34 (.34) | 1.47 (.45) | 2.32 | .02 * | 1.63 (.09) * |
| V7 | .03 (.04) | .02 (.04) | .60 | .28 | .08 (.12) | .23 (.27) | 4.73 | .00 * | .29 (.24) |
| V8 | .11 (.09) | .14 (.13) | - .93 | .37 | .01 (.02) | .01 (.02) | - .00 | 1.00 | .00 (.00) |

## Empirical validity (V1) and precision of concepts (V2)

As shown in figure 7.16, the current study and the study carried out by Vennix correspond with respect to the variables that refer to the degree to which the external model is used in the conceptualization of the problem: both Vennix (1990) and this study report that the differences between pre- and posttest are significant with respect to the variables V1 and V2. In addition to that, Vennix reports that significant differences are found for the strategic variables 'delays' and 'density' (connectivity) as well.

However, the degree to which the participants have incorporated the model, varies substantially between the two studies. In the Vennix study, on average, about 99 per cent of the concepts and 90 per cent of the relationships used in the policy notes written by the participants on the posttest are in accordance with the external model (Vennix, 1990, p. 199). In the present study however, only figures of about 45 and 60 per cent are found. One of the most important reasons for this difference probably is the fact that in the Vennix study participants were asked to write about something they already had relatively much knowledge about: they were asked to write about an economics problem of which a considerable amount of knowledge exists. In the present study however, participants were asked to write about a problem (the increase of costs of the Dutch Health Care system) for which hardly any well-established theory exist. The participants abviously have experience and knowledge of the health care system, but this knowledge is scattered among many people and there is no such thing as a 'proven' set of assumptions or statements. As a consequence, not only do the participants of the present

[145]Because we are concerned with the effects brought about by the participative policy modelling method, the outcomes of the study will be compared to the effects produced at the experimental group (Vennix, 1990) only. Moreover, because the experts participating to the Vennix study filled out a pretest only, the group that will be used to compare our findings to, will be the group of students who participated to the Vennix (1990) study. In addition to theses comparisons, some comparisons between the pretest scores of the present study and the pretest scores of the expert group will be made.

study have lower pretest scores to start with, it may also be that, because the external model is not regarded as a fixed and true (established) model, they are less likely to incorporate it in their conceptualization. This may account for the fact that the posttest scores of the present study are still considerably lower than the posttest scores found in the study carried out by Vennix. In addition to this, the Vennix scores may have been increased artificially because the coding procedure carried out in the Vennix study was much more aimed at translating the concepts (and consequently the relationships) into concepts (and hence relationships) used in the model than the coding procedure that was followed in the present study (Vennix, 1990, pp. 265-268).

Scope (V3)

Regarding the scope of the conceptualization (V3), defined as the proportion of concepts (the number of concepts divided by the number of lines), a substantial difference is found between the two studies. In the Vennix study, subjects wrote .43 new concepts per line, whereas the subjects of the present study wrote on average 1.32 new concepts per line. Since we do not have any figures on the number of concepts and lines for the experimental group, scores regarding the entire Vennix group (experimental and control group) will have to be used to account for the differences in scope between the two studies. In the Vennix study it was found that participants on average used 20.44 unique concepts on the pretest and 21.96 unique concepts on the posttest (Vennix, 1990, p. 189). In the present study, the number of concepts used on the pretest is equal to 24.78 on the pretest and 22.56 on the posttest (for the on-treatment group). Since no substantial differences are found between the two studies with respect to the number of unique concepts used by the participants, and both students and experts in the Vennix study have a substantially lower score than the participants of the present study have, one cannot argue that there must exist a difference in the number of lines that are used to describe those concepts in that students use more lines to write the same number of concepts than experts (policy makers) do. A more plausible explanation than the conciseness of one's writing style is that in the present study participants were asked to answer a particular question whereas the participants of the Vennix (1990) study were asked to write a policy note. It may well be that differences exist in the way in which people respond to a question and the way in which they express themselves when writing a policy note. Moreover, the ill-structuredness of the problem may be a factor that has affected the 'scope' score. Note that in the Vennix study participants have a much more integrated conceptualization - they refer to a concept used previously more often than those participating in the present study do. It may well be that when confronted with an ill-structured problem, people just use one and the same concept less often than people confronted with well-structured problems do, as a result of which they have a relatively high score on the 'scope' variable.

Detail: Length of maximal paths (V4), and integration (V6)

Another interesting difference concerns the average length of maximal paths (V4). Despite the fact that the participants in this study were dealing with an ill-structured problem, the average length of the maximal paths included in their cognitive maps is substantially higher than in the study carried out by Vennix. Note that this relatively high score on average length is not due to the fact that the cognitive maps of the participants are more connected, for the people participating in the present study scored lower on density (V6) than the students of the Vennix study did (probably due to the fact that the students were dealing with a relatively well-structured problem and the kind of model that was used - an economics model). As such, it suggests that policy makers, on average, use longer cause-effects chains in their discussion of the problem that novices (students) do. This conclusion is supported by the fact that in the Vennix study experts have longer maximal paths than student have (Vennix 1990, p. 199).

## Detail: Length of feedbackloops (V5)

Regarding the average length of feedbackloops included in one's conceptualization, note that both students and experts of the Vennix study have a higher score than the policy maker participating in the present study. One possible explanation for this stems from the artificial character that some of the feedbackloops may have. It is believed that artificial feedbackloops on average are longer than feedbackloops that are created deliberately,[146] [147] for longer feedbackloop are considered to be more difficult, as a result of which they are less likely to be incorporated deliberately.

## Time-factor (V7)

The percentage of time-indications (delays) is low in both studies, although students do have a substantially higher proportion of delays included in their conceptualization than the policy makers participating in the present study do. This difference may be due to the difference in topic that is dealt with in both studies: in the Vennix study, participants were confronted with an economics problem, whereas the people participating in the present study were asked to comment on public health issues. It may be that delays (i.e. time-indications) are more common sense in economics (dealing with time-series for instance) than they are in public health. Another possible reason for this difference is that only when certain levels of domain-specific knowledge have been arrived at, additional refinements in one's conceptualization (e.g. use of time-indications) can be expected to be made (although the number of concepts and relationships used by the participants does not differ substantially between both studies, the degree to which elements from the external model have been used differs significantly, showing that indeed a difference may exist in the degree to which the participants feel the knowledge they have about the problem can be looked upon as relatively 'valid' as a result of which they can start concentrating more on non-content related aspects such as time-indications). Another indication for the fact that those participating in the present study may feel differently about the status of the knowledge they have about the problem at hand than those who took part in the Vennix study did, is the large percentage of new concepts and relationships that the present study's participants use on the posttest (more than 50 percent of the concepts ad relationships used by the participants had not been used on the pretest). In addition to this, note that differences in the programs offered to the participants may have contributed to the difference between the two studies with respect to the number of time-indications included in the participants' conceptualization of the problem as well. Not only did the Vennix study have more sessions as a result of which more profound learning effects may have been brought about, one particular session in the Vennix program was devoted to delayed relationships (Vennix, 1990, pp. 92-93), whereas no special attention was given to the issue of time-indications (delayed relationships) in the present study's participative policy modelling program at all.

---

[146] As mentioned previously, it is believed that a relationship between connectedness and number of feedbackloops exists in that the more connected a cognitive map is, the more likely it is that the feedbackloops found in it are of an artificial nature. This belief is to some extent supported by the Vennix study for it shows that (compared to the outcomes of the present study) a higher score on connectedness is found together with a relatively high score on the number of feedbackloops. However, note that the number of feedbackloops included in the Vennix model and the amount of time spent on feedbackloops may account for this difference between the Vennix (1990) study and the present study as well.

[147] Calculations carried out by Vennix, revealed that the econometric model used in the Vennix (1990) study contained about 10,000 possible feedbackloops, whereas 'only' 700 possible feedbackloops were detected in a system dynamics version of the very same model, while both models contained about the same number of endogenous variables (i.e. 40).

### Societal conditions (V8)

Finally, with respect to the proportion of exogenous variables included in one's conceptualization (V8), participants in the present study do much better than the students in the Vennix study. However, one should take into account that differences exist in the way in which exogenous concepts have been defined in both studies. In the present study, concepts qualify as exogenous if they either refer to something completely outside the model, or refer to another sub-system than the sub-system the question was referring to. In the Vennix study however, concepts only qualified as exogenous concepts if they were able to meet the first requirement. Had the very same definition of 'exogeneity' been used, the difference between the two studies with respect to V8 would not have occurred, for in the present study hardly any reference is made to something completely outside the model as well. As such, a close correspondence exists between the findings of the present study and the results presented by Vennix (Vennix, 1990, p. 193,199).

Having compared the student-group taking part in the Vennix (1990) study with the policy makers taking part in the present study, some comparisons between the expert-group (Vennix 1990, p. 199) and the policy makers of the present study will be made. Since the differences between the students-group and the present study's policy makers have already been discussed, account will only be taken of those variables where significant differences between the experts and students of the Vennix study exist. Consequently, both groups of experts will be compared with respect to the variables *'model concepts'* and *'density'* only[148].

Regarding the proportion of *model concepts* used on the pretest, it is shown in figure 7.16 that experts already think in terms of the established theory (that is, the external model) prior to taking part in the sessions: 93 per cent of the concepts they use to refer to the problem are concepts that are used in the external model as well. The experts participating in the present study however, have no established theory to rely on, as a result of which their correspondence to the external model is much lower (in fact, about 58 per cent of the concepts they use come from the external model).

As far as the differences in *density* between the two groups of experts are concerned, the Vennix group clearly outperforms the participants of the present study. Participants of the present study have a much less integrated cognitive map than the experts of the Vennix study have.

To sum up the above comparison of the present study and the study carried out by Vennix (1990), an increase in the percentage of concepts and relationships used from the external model was reported in both studies, indicating that participative policy modelling does indeed lead to an increase in the use of the external model used, discussed and adjusted during the model building sessions.

Comparisons between the two studies, however, also revealed differences that probably stem from the different kind of model (or kind of problem) that is used in both studies.

It is our belief for instance, that because the Vennix study focused upon a moderately or well-structured problem, in contrast to the present study where participants were asked to think about an ill-structured problem, participants in the present study have lower absolute scores on variables related to the use of an external model (V1 and V2) than participants in the Vennix study have (this applies to both students and experts from the Vennix study).

Moreover, it is felt that, although no exact figures are known for the Vennix study, the number of concepts and relationships that are newly used on the present study's posttest, also serve to illustrate the ill-structured character of the health care issues dealt with in the present study. Besides, it shows that those taking part in the present study apparently had difficulty conceptualizing the problem, and were willing to change about

[148] In the Vennix study, experts use significantly more concepts of the external model on the pretest than students do, and have a significantly higher cognitive map density than student have.

50 % of their pretest conceptualization at the time the posttest was taken. The difference in the level of structuredness of the problem focused upon in the sessions is also reflected in the difference that is found in the integration scores reported by both studies. Differences that exist despite the fact that the average total number of concepts and relationships is about the same in both studies.

Comparisons between the two studies thus show that in both cases (studies), participants have made an effort to incorporate elements of the external model focused upon during the sessions. In addition to this, it has shown that the absolute value of the pre- and posttest scores seems to be considerably affected by the structuredness of the policy problem focused upon.

## 7.7 SUMMARY

In the present chapter, an overview was given of the outcomes of this study. A general description of the cognitive maps of those taking part in the sessions was given first. This description showed that no increase in size was found due to taking part in the sessions. It also showed that the cognitive maps of those participating in the present study are of about the same size of those taking part in the Vennix (1990) study. In addition to that, it was found that some 50 per cent of the concepts and relationships used on the posttest were new compared to the concepts and relationships used on the pretest.

With respect to the five major research questions focused upon in the present study, no significant change or improvement was found, with the exception of the fifth question concerning the degree to which participants are aware of each others' point of view. However, note that this change was assessed by means of the questionnaire only.

In discussing the reasons why the outcomes differed from what we had expected, it was argued that from an observer's or outsider's point of view we probably would have to conclude on the basis of the present study's outcomes that participative policy modelling is not able to bring about a change in conceptualization. However, from an insider's point of view (i.e. by moving from a Trivial Machine approach to a Non- trivial machine approach), opening the black box, we see that all sorts of processes have played a role, context-sensitive processes, that may have affected the usefulness, meaning and incorporation of the program offered.

Following the discussion of the importance of such a change in perspective in order to be able to understand the (lack of) effects found in the present study, the influence of the potentially specifying variables on the relationship between participative policy modelling (treatment) and conceptualization was examined. Prior to this however, an overview was given of the potentially specifying variables themselves rather than their influence on the relationship between treatment and effect.

As far as the evaluation of the program is concerned, it was found that participants on the whole did appraciate the small-group activities most and the workbook least. The found the program rather interesting and useful, though they claim not to have learned from it too much. The on-treatment group spent on average 14.58 (sd = 2.58) hours on the program, which is much less than the 22 hours and 55 minutes that were spent on average by the participants of the Vennix study.

Regarding the specifying effects of the above-mentioned variables, educational background, years with the firm, department, amount of time spent, and age proved to have a differentiating effect on the relationship between individual change and the program offered. Moreover, these variables also proved to have a specifying effect on the way in which the modelling processes itself was perceived.

With respect to the specifying effects of the variables that are concerned with the evaluation of the program, the program's perceived length, its usefulness and the amount of learning that participants attributed to the program had a specifying effect of the amount of individual change that had taken place. However, as with all of the above potentially specifying variables, no effects on inter-individual change and changes in the awareness of the others' point of view could be concluded to.

Finally, comparisons with the Vennix study were made to put our findings into perspective. In order to be able to do so, some new variables had to be constructed. Although some striking similarities between the two studies were found (e.g. concerning the increase in use of the number of elements of the external model (preliminary model)), it is important to bear in mind that differences in the actual program, the subjects taking part in it, and the kind of model employed, make it difficult to draw general conclusions on the basis of these findings.

# CHAPTER 8: REFLECTION

## 8.1 INTRODUCTION

Now that the outcomes have been presented, the implications of these findings for the assumed relationship between participative policy modelling and policy making in general and the relationship between participative policy modelling and the conceptualization of a problem in particular will be focused upon. To start with, a brief overview will be given of the reasons why the present study was carried out, how it was carried out, and the major outcomes that were found.

In discussing the implications these outcomes have on the above-mentioned relationship between participative policy modelling and policy making, a distinction will be made between so-called narrow implications, that is, implications within the approach taken in the present study, and broad implications, i.e. implications that supercede or question the approach taken in this study. As part of the latter, the present study's preoccupation with cognitively-based support will be questioned, resulting in the presentation of an alternative framework for understanding and examining the effects of participative policy modelling. The alternative framework serves to provide a possible explanation for the somewhat unexpected (empirically-based) outcomes of the study. Note that the framework should not be considered as a 'proven' frame of reference. However, it is meant to serve as a first step in the process of reconsidering contributing to the policy making process (for instance by having policy makers participate in the policy making process). As such, it aims to meet the desire to question the present study's framework itself as expressed in Chapter One. Having discussed this possible alternative frame of reference, work is discussion, some of the limitations of our study will be focused upon and recommendations will be made for studies to come. These recommendations concern both the evaluative character of the study (how future evaluations should be carried out) and the way in which participative policy modelling programs should be designed in future, to become more effective. Finally, some practical implications of the present study will be presented.

## 8.2 EVALUATING THE EFFECTS OF PARTICIPATIVE POLICY MODELLING

In the current study, the question as to whether participative policy modelling is able to bring about changes in the way in which those who participate in it look upon a particular policy problem has been focused upon. It was decided to focus on the participative policy modelling's potential to cognitively support the policy making process because the conceptualization of the problem can be regarded as one of the most important stages of the policy making process when dealing with ill-structured problems (Brewer & deLeon, 1983; Dunn, 1981; Gusfield, 1981; Kalff, 1989; Mason & Mitroff, 1981; Quade, 1989).

Although the successfulness of participative policy modelling as a tool to support the policy making process has been reported on several occasions (e.g. Akkermans, 1992; Kim, 1989; Lane, 1992a, 1992b; Larsen, Morecroft, and Murphy, 1991; Morecroft, 1992; Richardson, Andersen, Rohrbaugh, and Steinhurst, 1992), hardly any empirical evidence for this support has been given so far, with the exception of the study carried out by Vennix (1990). Hence it was decided to use that study as a frame of reference, that is, adopt those elements of the study that proved to be very useful, and change the elements where it was felt that room for improvement existed[149] [150]. Moreover, using the

---

[149] Most of the guidelines concerning the actual participative policy modelling program, for instance, were taken into account when designing the present study' s program. A more detailed description of the way in which these recommendation were built into the present study's program will be given in the section in

Vennix (1990) study as a frame of reference not only enabled us to make use of the knowledge acquired in it (his experiences in applying a particular methodology), it also allowed us to put the present study's findings in a broader perspective by relating the outcomes of the present study to the outcomes of his study (outcomes that concern both the methodology taken and the study's findings (facts)).

To examine the effects of the participative policy modelling program on the participants' conceptualization of a policy problem, a distinction was made between the ***domain-specific*** aspects and ***strategic*** aspects of a person's conceptualization (Klabbers, 1990). The distinction between domain-specific knowledge and strategic knowledge was introduced to be in the position to examine whether participative policy modelling is able to change those aspects of the participants conceptualization that are related to the specific content that is being dealt with (e.g. health care, oil-industry, personnel planning), or those aspects that are less dependent on a particular application and seem more related to the way in which domain-specific knowledge is being organized or applied to arrive at a solution for the problem at hand.

In addition to this, a distinction was made between ***individual*** and ***inter-individual*** changes in conceptualization. This because it was believed that policy making is often carried out by groups of policy makers rather than by individual policy makers only (Isaacs & Senge, 1992; Hall, 1984; Reinhardt & Schwelker, 1992; Scheper, 1991; Vennix, 1990; Weick & Bougon, 1986;). Combining the domain-specific and strategic knowledge dimensions with the distinction between individual and inter-individual changes, four basic research questions were arrived at, all concerning a change in the participants' conceptualization of the problem at hand.

With respect to the changes in the way in which ***individual*** participants look upon the problem, it was expected that participative policy modelling was able to improve both the domain-specific and strategic aspects of the participants' conceptualization of the problem. Regarding the ***inter-individual*** change in conceptualization, it was expected that taking part in the participative policy modelling program would result in a decrease in variance among the participants' domain-specific and strategic conceptualization of the problem. In other words, participative policy modelling was expected to be able to bring about a shared understanding among those who participated in it.

Besides the changes brought about in the knowledge participants have about the problem at hand, the effects of the program on the knowledge participants have about ***each others' conceptualization*** of the problem at hand were examined as well. Not only because it was suggested by Vennix that communication is *"an important aspect of computer modelling that warrants the time and effort of systematic empirical research"* (Vennix, 1990, p.220), but also because increased knowledge of the other participants' point of view can be regarded as a prerequisite for shared understanding[151]. It was expected that participative policy modelling would be able to increase the participants' knowledge of each others' conceptualization, because of the interaction that is taking place between the participants (particularly in the small-group activities).

To arrive at criteria to assess the degree to which a change in domain-specific and strategic knowledge is brought about by the participative policy modelling method, the ***system dynamics framework*** was used. This because participative policy modelling can

which the present study's recommendations concerning the design of participative policy modelling sessions will be presented.

150 For a detailed description of the differences between the present study and the study carried out by Vennix (1990), the reader is referred to Figure 2.15 in Chapter Two.

151 This is not to say shared understanding cannot be brought about without intensive communication between all participants. It may well be possible that a shared cognitive map is arrived at by simply listening to a very charismatic leader, or by reading one and the same book. However, in the context of being supportive to the process of tackling ill-structured policy problems, group-think alike shared understanding seems less appropriate than intensive communication among the participants to arrive at commonality or shared understanding.

be considered as a participative version of system dynamics modelling - most of the activities carried out in the participative policy modelling program correspond to activities carried out in system dynamics modelling.

To evaluate the effects of participative policy modelling, an untreated control group design with pretest and posttest (Campbell & Stanley, 1963, Cook & Campbell, 1979) was selected. However, due to reasons beyond our control (restrictions posed upon the present study by the participating organization), some changes had to be made to this research design as a result of which the design actually used in the present study is the so-called ***One-Group Pretest Posttest Design*** (Campbell & Stanley, 1963, Cook & Campbell, 1979).

The program offered to the participants consisted of ***four sessions***. Prior to each of the sessions (with the exception of the fourth), participants were asked to read a preparatory text called 'workbook'. During the sessions, activities were carried out in both a small-group and plenary mode. To speed up the process of model building (only three sessions could be spent on building and playing with the system dynamics model), use was made of a ***preliminary model***, that had already been built by the researcher.

Since no direct access can be made to the way in which participants look upon a particular policy problem, participants were asked to fill out a ***pre- and a posttest*** containing three questions about the effects of policy measures that could be taken to reduce the costs of health care. In addition to this, a ***questionnaire*** was completed by the participants at the end of each session to evaluate the program, to assess whether taking part in the program had made them change their conceptualization, and to measure whether they felt the program had given them the opportunity to communicate with each other. The questionnaire was also used to assess whether a relationship existed between the way in which participants evaluate the program and the effects the program has had on their conceptualization (the potentially specifying variables were used for this).

To transform the answers given to both pre- and posttest into data suitable for analysis, it was decided to follow Vennix (1990) and make use of the ***cognitive mapping*** approach (Axelrod, 1976, Bonham, Shapiro, and Heradstveit, 1988; Eden, Smithin & Wiltshire, 1980; Klein & Cooper, 1982; Vennix, 1990; Weick, 1979; Weick & Bougon, 1986). This because the cognitive mapping approach allows one to recode written statements in such a way that only (analyzable) causal relationships between concepts remain[152]. Based on these so-called cognitive maps, representing how individual participants look upon a particular problem, values for the criteria that stem from the system dynamics modelling approach were arrived at, so that pre- and posttest scores could be compared.

## 8.3 OUTCOMES AND IMPLICATIONS

Having presented the objectives, background and design of the present study (*what* has been evaluated, *why* it has been evaluated, and *how* it has been evaluated), a brief summary of the most important findings of the present study will be given to be in the position to discuss whether participative policy modelling is indeed capable of living up to its expectations - supporting policy makers in the process of policy making. As part of this discussion, some of the present study's limitations will be focused upon, for these limitations clearly affect the conclusions that can be drawn from the present study. Note that the discussion of the implications of the outcomes is divided into two parts. In the first part, the implications of the outcomes of the present study for the relationship between participative policy modelling and policy making is considered from what can be called a 'traditional rationalistic' approach according to which the effects of participative policy modelling can be assessed on the basis of pre- and posttest measures

[152] Hence, such cognitive maps are often referred to as cause maps (Weick & Bougon, 1986, pp. 106-107).

provided the study has been carried out scientifically[153]. Consequently, in discussing the implications of the present study's findings on the (claimed) relationship between participative policy modelling and the conceptualization of policy problems, the present study's scientific shortcomings will be presented in detail and recommendations for future studies will be made. Following this discussion, the implications of the present study's outcomes will be examined from a more radical point of view in that the basic premises of the 'traditional rationalistic' approach taken in the present study will be questioned, resulting in a critical review of the present study's 'what' rather than the scientific status of its 'how'. Following this review, some conclusions with respect to the design of participative policy modelling programs to be will be drawn, based on the outcomes of the participant-based evaluation of the participative policy modelling program.

## 8.3.1 IMPLICATIONS FROM A RATIONALISTIC POINT OF VIEW

Since the outcomes of the study have already been described in detail in Chapter 7, only a brief summary of the outcomes will be given in the present chapter. This to be in the position to examine the implications of these outcomes for the claim that participative policy modelling is able to bring about a change in the way in which participants look upon a policy problem and thus contribute to the policy making process.

As discussed in the previous chapter, no significant changes in conceptualization were found neither with respect to the domain-specific and strategic dimension, nor related to the individual and inter-individual dimension. Some change was found in the knowledge participants have about the way in which other participants look upon the problem. As far as the evaluation of the program was concerned however, most of the participants expressed they had enjoyed the program in that they had found it interesting, had learned from it, had found the amount of time spent on it about right, and had found the program useful.

Despite those relatively positive participant-based evaluations, the pre- and posttest measures clearly suggested that participative policy modelling as presented by Vennix 1990 and in applied in the present study as well, is not able to live up to its expectations, that is, bring about a change in the way in which those who take part in it ***conceptualize*** the problem, for no significant differences between post- and pretest were found. In the previous chapter, it already was suggested that changes in the present study's context may have been responsible for these (lack of) outcomes. However, prior to concluding that either participative policy modelling is not able to bring about the expected change in conceptualization, or suggesting that the effects looked for should be different and consequently the approach taken as well because of the context-sensitivity of knowledge in the policy making process or the nature of the policy making process itself, account should be taken of the present study's limitations for they may have contributed to the lack of outcomes to start with. So rather than questioning the approach taken - its rationalistic paradigm - in this section by introducing and discussing an alternative frame of reference, some of the most important limitations of the present study will be considered first, for they may also be held responsible for the somewhat surprising (lack of) results.

### Limitations

The present study is limited in several ways. First, as already stressed in Chapter 1 (figure 1.1), the policy making support that is focused upon is of a conceptual, cognitive nature. Not only because the conceptualization of the policy problem is considered to be one of

---

153 The 'traditional rationalistic' approach was discussed in section 7.3.6 under the heading of the Trivial Machine approach. As described in section 7.3.6, to assess the effects of a particular treatment from a Trivial Machine point of view, use is made of correlation between in- and output pairs primarily.

the most important phases of the policy making process when dealing with ill-structured problems, but also because assessing the effects changes in conceptualization are having on policy choices made by policy makers and the influence these decisions eventually have on the performance of the organization, are extremely difficult to assess due to the influence of disturbing and confounding variables (Cowan, 1992; Schaik, 1988; Vennix, 1990). Moreover, one should bear in mind that changing one's conceptualization not necessarily implies that differences in policy making behaviour are brought about (it may well be that people just become more confident in the decisions they make, without changing them), or that differences in organizational performance will be established (Reimann & Ramanujam, 1992).

Second, the absence of a control group obviously affects the internal and external validity although it has been argued that most of the potential treats to internal validity could be made implausible. As far as the external validity is concerned, the outcomes cannot and should not be generalized to other situations, other groups of participants and other periods of time. As such the knowledge acquired through the presents study suffers from the very same situationality or context-sensitivity as the knowledge acquired in participative policy modelling sessions does (cf. section 7.4).

Third, there is no frame of reference to compare the effects of participative policy modelling to. This is partly due to the absence of a control group described above, but also because no comparisons have been made between the effects brought about by the participative policy modelling method and other policy support methods. As such our study follows the limitation described by Bredemeyer and Greenblatt's (1981) *"So how's your wife Sam?, Compared to what?"*. The only study that can be used as a potential frame of reference to put our findings into perspective, is the one carried out by Vennix (1990).

Fourth, as far as the present study's design is concerned, the study is limited in that the posttest was conducted within 2 weeks from the last session whereas changes in conceptualization may require more time to materialize in one's conceptualization of the problem: *"the study of real-time thinking must take into account potential time lags for anticipated changes"* (Cowan, 1992, p. 396). Moreover, the changes in the domain-specific and strategic dimensions of one's conceptualization have been measured at the same time. Breuer and Kummer (1990) however, believe that more time is required for strategic knowledge to be incorporated in one's thinking than domain-specific knowledge. In addition to this, one can question whether strategic and domain-specific knowledge should be measured using one and the same instrument (open-ended questions). Since the questions given to the participants focused on the costs of health care in a particular section of the health care system, it is possible that too much attention is paid to the content of one's representation and that too little attention is given to its strategic aspects (the way in which this domain-specific knowledge is organized). Hence an approach in which new problem area's (new domain-specific problems) are given to the participants (Bakken, 1989; Gould, 1989b), thus focusing on the transferability of (strategic) knowledge, may be a more appropriate way of assessing changes in the use of system dynamics principles (i.e. strategic knowledge) to deal with complex, ill-structured problems.

Fifth, regarding the distinction between individual and inter-individual changes in conceptualization, the present study is limited in that both have been measured in the same way, using the same measurements. Future studies may consider using separate measures for individual and inter-individual conceptualizations for instance by asking the participants as a group to write a policy document.

Another limitation of the present study related to the way in which the changes in conceptualization have been assessed, is the lack of motivation that was found among many of the participants when they were asked to answer the very same questions twice (on both the pretest and the posttest). The consequent lack of motivation may have affected the quality of the posttest measures, and hence have reduced the likelihood that changes were found between post- and pretest. To account for this lack of motivation, bear in mind that answering the questions on the pretest had already taken a considerable amount of time and that the participants were under a lot of (time) pressure due to changes in the environment of the organization and changes within the organization itself.

Seventh, one can question whether the right program was given to the participants. Not only because the number of sessions and the amount of time invested in the program seem rather little to achieve any major changes in conceptualization, but also because no special attention was paid to system dynamics characteristics in action[154], despite the fact that acquiring a system dynamics point of view had been considered an important objective for the organization taking part in the program. Hence, some discrepancy may have existed between the objectives of the participative policy modelling program and the curriculum the participants took part in.

Another limitation of the present study's participative policy modelling program is that use had to be made of a preliminary model to speed up the modelling program. Although hardly any adjustments were made to this preliminary model by the participants, this should not be regarded as an indication of the preliminary model's success as is done by for instance Kim (1989a, p. 331), for it is possible that participants, because of lack of ownership (they did not create the model themselves, the model was created by us, the expert modellers), simply passively accept the model and its conclusions without questioning it and revealing or changing their own assumptions (Morecroft, 1992).

Ninth, although the use of the cognitive mapping approach seemed very promising to start with (many instances have been reported where use was made of the cognitive mapping approach), only limited empirical knowledge is available regarding the size and structure of cognitive maps and how they develop over time. Hence it is difficult to put the findings of the present study into perspective - we do not know how easily cognitive maps can be changed, how stable they are over time. Moreover, the cognitive mapping approach has proved to be rather time-consuming. A lot of time and effort is required to transform the texts written by the participants into cognitive maps. Besides, several coders are needed to assist in the process of coding the texts in order to be able to arrive at reliability scores. However, these persons need to have both knowledge of and be skilled in the procedures to transform the texts into analyzable data. For this, it is important that they also have good working knowledge of the health care system. Because people who meet both requirements and are willing (have the time) to assist in the process of coding, are hard to find, it seems worthwhile to consider different (less time-consuming) approaches to arrive at cognitive maps[155].

Another limitation of the use of cognitive maps is that we cannot be sure whether cognitive maps are capable of differentiating between more an less sophisticated conceptualizations. Although differences in scores on criteria such as the number of concepts and relationships can easily be arrived at, it is questionable whether those aspects suffice to represent complex objects such as one's conceptualization of a problem (Scheper, 1991). To illustrate this limitation of the use of cognitive maps, it is possible that having given it a good thought, participants decide not to change anything in their conceptualization because they now know for sure (thanks to taking part in the program) that their conceptualization is an appropriate one. Since the no changes are found in the elements included in one's conceptualization, such a change in the status of the elements

154 Some information on the major elements of system dynamics modelling was given in the first workbook. However, the effects of delays and feedback mechanisms were not discussed and focused upon to provide the participants a better understanding of the system dynamics principles underlying the dynamic behaviour of the model.

155 One could for instance ask participants to draw their own cognitive map as part of the pre- and posttest thereby reducing the amount of time spent by the researcher on the coding process and eliminating the need for other coders to be able to arrive at reliability scores.

included in the conceptualization is not reflected in any way in the scores participants are given on the basis of the cognitive mapping approach.

Eleventh, the question arises whether the right organization and the right people took part in the research. As far as the organization taking part in the program is concerned, the dramatic changes the organization was confronted with at the time of the sessions (post-merger problems, major changes in the health care system, new responsibilities and tasks (e.g. budgeting)), clearly have affected the organization's ability to invest in the program negatively. These dramatic changes also resulted in the CEO's inability to attend the sessions (he only attended part of the second session), which may have affected the participants perception of the importance of the program.[156 157].

Regarding the people taking part in the program, two comments need to be made. Firstly, it is our impression that participative policy modelling should be carried out with policy makers rather than executive decision-makers (Vickers, 1965)[158], because it is only policy makers who do actively (re-) define and reconstruct the reality they work in. Executive decision makers by contrast, are more likely to operate within the reality (scope, definition etc.) offered to them by for instance policy makers, as a result of which they do not seem the people who benefit most from taking part in a program in which the reality they live and work in is being defined (and hence constructed). It is important to note with respect to this distinction, that not all the people attending the sessions can be considered as policy makers in the strict sense. Some of the participants seem to be executive decision-makers rather than policy makers. Secondly, the present study is limited in that only 22 participants filled out a pre- and posttest, of which only 18 did attend to such an extent that they could be regarded as true participants, that is, as people belonging to the on-treatment group. In order to find any significant differences in such relatively small groups, the differences between pre- and posttest will have to be relatively large.

## 8.3.2 IMPLICATIONS FROM A NON-TRIVIAL MACHINE POINT OF VIEW

Because of the limitations described in the previous section, one cannot and should not conclude that participative policy modelling is **unable** to bring about a change in conceptualization despite the fact that no significant differences in pre- and posttest were found. One of the conclusions of this observation could be that future studies should be

---

156 It is interesting to see that when time is needed to deal with dramatic changes in order to respond in a satisfying manner, only little time seems available and one is often rushed to act rather than to think first, thereby jumping at conclusions One way to avoid this problem, is to consider using management support tools such as participative policy modelling in anticipation of changes to come (De Geus, 1988, Kalff, 1989), rather than waiting till there is hardly any time available to considered suitable actions

157 At one stage, the researcher overheard a conversation three members of a project team were having in discussing the effects of governmental initiatives regarding ethical issues As one of the team members asked the others whether the initiative would have a positive effect on the use of health care and hence the costs, one of the other team members commented that because government had come up with this initiative, it had to have positive effects and that they should not question this for they had only a limited amount of time to consider the consequences of this initiative for their own organization This clearly illustrated the pressure the organization was under at the time of the program, and the need that was felt to act rather than to think This is not to say that the members of the organization were not capable of considering the initiative carefully, but that they simply felt they could not afford spending much time thinking about it Hence, it should not come as a surprise that some participants also felt they were not in the position to spend much time on the participative policy modelling program, the preparatory texts, and the 'time-consuming' posttest

158 The very same difference also seem to hold for program leaders and program managers. It can be argued that program leaders are those who operate within a particular definition of reality (a pre-defined scope), whereas program managers need to be actively involved in setting the program's scope and defining the problem to be solved.

better planned and designed (e.g. by improving the design, increasing the degree of participation, and increasing the number of people that take part in it) so that the detection of differences in conceptualization would become more likely, as a result of which one would be in the position to conclude that participative policy modelling is indeed able to enrich the conceptualization of those participating in it.

In the present section however, the approach taken in the study itself will be reviewed critically and an alternative theoretical framework will be introduced to increase our understanding of why no significant effects were found. This alternative framework may also assist in the design of participative policy making studies to be.

According to this alternative, the approach taken in the present study can be questioned for two reasons - for not taking into account the situationality of the knowledge being used in the policy making process, and for not focussing as much as should be done on the process character of policy making when evaluating the effects of participative policy modelling. It is our impression that the possibility of an alternative theoretical framework has only come to mind now that participative policy modelling has been applied and evaluated in a real-life situation with real-life policy makers, for *"without capturing the real-time characteristic of executive cognition, much of the richness of the phenomenon may be lost"* (Cowan, 1992, p. 396).

Applying participative policy modelling to and evaluating it in a real-life context (a real organization and real-life policy makers rather than students in a laboratory setting in which much of the context can be filtered out), has made us realize that because hardly any changes were found, the present study's theoretical basis, its 'what', may have to be reconsidered critically. The first aspect of the study's 'what' to be reconsidered, concerns its preoccupation with cognitive changes and the notion of knowledge that is being used for this. Secondly, to account for the somewhat surprising results attention will be paid to the way in which the policy making process has been defined in the present study. It is expected that by concentrating more on the *process* side of policy making, it can be made plausible why it is difficult to find any effects in product- or output-based evaluations.

In section 8.3.1, it was argued that if we would look upon the effectiveness of the participative policy modelling process from an outsider's perspective[159] and look upon the social system where the participative policy modelling program was carried out as a so-called Trivial Machine, using correlations between in- and output pairs (pre- and posttest scores) to assess whether a change in conceptualization is brought about, we would probably conclude that participative policy modelling is not be able to contribute to the policy making process.

The present study is not alone in its conclusion that it is difficult to get other people to understand the dynamics of (complex) system dynamics models. The difficulty of getting other people understand the structure and behaviour system dynamics models have been reported in literature as well (Anderson, 1970; Boocock & Schild, 1968; Doerner, 1980, Doerner, Kreuzig, Reither, and Stäudel 1983[160]; Klabbers, 1972[161];

---

[159] An outsider's perspective is a perspective where the researcher is not being a part of the system it is observing or examining. The system under consideration is considered as a Trivial Machine, that is, a machine that is behaving in a completely predictable manner. The findings that concern the system do not concern the researcher him- or herself. The outsider's perspective is, as will be explained in more detail in secxtion 8.4, in opposition to the insider's perspective.

[160] According to Doerner (1983), no clear effects of a knowledge-based support of managers in their policy making process have been found.

[161] Klabbers (1972) for instance, found that even after six months of intensive training with a relatively simple model, participants occasionally had difficulties predicting and understanding the dynamic behaviour of the system at hand. Hence using participative policy modelling as a tool to convey understanding / knowledge of a relatively complex model such as the model of the Dutch health care system in only three sessions, seems a hard task to accomplish.

Vennix, 1990[162]). One of the issues that system dynamicists have been focused upon over the last two decades, is the issue of utilization of knowledge: how to make other people (policy makers in particular) employ the knowledge than can be acquired by being involved in a model building process. Although it has been recommended to reduce the size of the models and to use the model building procedure for ill-structured problems, the most important recommendation following this discussion concerned the need for client participation in the model building process: *"From these recommendations the most important seems to us the need for more client involvement in the modelling efforts. At the same time, however, this is also the most radical and probably the one that is most difficult to realize* (Vennix, 1990, p. 36). In the present section the importance of client participation will not be questioned, however, it will be argued that there are additional reasons for having policy makers participate in the process.

Because participation has not been able to resolve the problem of knowledge utilization altogether (Kalff, 1989[163]; Vennix, 1990), as illustrated by the present study as well, it is suggested to question the relationship between participative policy modelling and supporting the conceptualization of those taking part in it.

The first step on this route was taken in section 7.3.6, where it was argued that if one would look at the relationship between participative policy modelling and cognitive changes from a different point of view, an insider's rather than an observer's point of view, a different (richer) perspective on the phenomenon under consideration (i.e. conceptual change) could be acquired.

> "To view evaluation through the traditional scientific input/output research paradigm alone presents dangers. The multiple and complex nature of inputs, the difficulty of detecting and measuring outputs, the essential impracticality of isolating outputs, the considerable and intrinsic interaction amongst participants, all pose problems" (Barnett, 1984, pp. 169-170).

Or as Cowan (1992, p. 396) put it:

> "It is possible that existing conceptualizations of executive thinking reflect social-epistemological assumptions (Greeno, 1989), that constrain potential understanding".

### *Situationality*

As a result of changing one's perspective from an outsider's to an insider's point of view, it was argued in section 7.3.6 that, as far as the present study is concerned, processes such as a changing environment (new regulations, dramatic changes on the health care market) and a changing organization (a merger that had just been completed and two departments which were in the process of being put together at the time of the present study), may have affected the meaning and usefulness of the program offered to the participants. Because of this, the knowledge offered to the participants in the form of the preliminary model may not have been incorporated as much as would have been the case if no such changes had taken place.

Based on this observation it was concluded that the knowledge provided through the participative policy modelling sessions is of a context-dependent or situational nature in that changes in context affect the meaning, applicability and usefulness of knowledge

---

162 Vennix (1990), had difficulty finding any major differences between pre- and posttest measures using the cognitive mapping approach as well.

163 Kalff's description of the participative policy modelling process carried out at Shell clearly indicates that other elements than knowledge play an important part in the success of the project.

offered to participants. A conclusion supported by Brown, Collons & Duguid (1989), Greeno (1989), and Cowan (1992).

The lack of effects found in the present study thus may have been caused by a change in the context of the sessions, as a result of which knowledge that seemed useful at the start of the project (the preliminary model) may have become useless or of minor importance. This would also account for the limited number of changes that were made to the preliminary model in the present study, for it was felt that in the present study the participants were *"passively accepting the model conclusions without revealing or changing their own assumptions"* (Morecroft, 1992, p. 26) rather than actively agreeing with the model.

Recently, system dynamicists have become aware of the context-sensitivity of the knowledge created and used during the sessions as well. Morecroft (1992, p.25) for instance, stresses the influence individual and inter-individual processes taking place during the sessions are having on the outcomes of the sessions:

> "It is easy enough for modellers to become fascinated with representation scheme's, software, the mental models of clients, and cognitive imperfections in dynamic decision making, and yet lose sight of individual and group behaviour on which models depend for both ideas and legitimacy."

System dynamicists have become aware that participative policy modelling cannot and should not be considered as an input to a client system which will automatically lead to predictable outcomes (whether those are learning effects or topics that should be covered during the sessions). That the social systems taking part in a participative policy modelling program have a will of their own and cannot be trivialized by us, is shown for instance by the role the CEO played in the second session of the present study's program - the CEO managed to change the topic under consideration at the small-group session he was attending completely. Similar experiences have been described by for instance Lane (1992, p. 341), who even uses Harold Wilson's "a week is long in politics", to illustrate the unpredictability and changeability of client organizations in participative policy modelling project. In other words, client systems cannot and should not be trivialized, they do not receive a program passively, but interpret, (mis)use and adjust the program according to their needs. Hence in developing a particular program one should not *"act unilaterally in the design of the learning environment, giving little freedom to participants"* (Isaacs & Senge, 1992, p. 191).

### *Process- versus product-oriented*

Looking differently at the policy making process carried out by groups of individuals, not only has made us realize that the knowledge used in the process is situational, depending on the people and situation participants are in, and the context in which the knowledge is used, it also enabled us to see that the policy making process itself may have to be defined differently, in that less attention should be paid to its so-called products (objectives, results), and more attention should be given to the process side of it. By changing our perspective, we move from being concerned about the 'premature closure' (jumping at conclusions) character of policy making to being concerned with the 'no closure' feature of the policy making process: *"they [ill-structured problems] have no closure. For ever evolving organization structures are an example in cases, where the implementation of a revised structure already sows the seeds for the next round of restructuring"* (Kalff, 1989, p. 53).

Policy making should be considered as a process-oriented rather than a product-oriented activity: it has no closure, there is no pre-defined objective to be reached for both the goal and the way to go are defined and continuously changed in the process itself[164]. As a consequence, the support that should be offered should not be concerned with (cognitively-based) products to be delivered only, but understand the role such products play in the process of policy making as well.

According to the Non-Trivial Machine metaphor, policy making support thus will have to take into account the contextuality of the knowledge provided by for instance participative policy modelling sessions, and be aware of the fact that the process-side of the intervention should be focused upon: "*...valuable insights about a policy problem are not so much the results of the final model. Rather these are obtained in the process of modelling...*" (Vennix, 1990, p. 40). A view supported by other contemporary system dynamicists as well (De Geus, 1988; Lane, 1992b)[165].

As a consequence, it is suggested that the lack of outcomes found in the present study may have to be attributed to the dilemma of evaluating a process (the contribution of participative policy modelling to the process of policy making) using a product-oriented approach.

Although an extensive philosophical and/or methodological discussion of the problem of assessing a process by means of a product-based evaluation would fall beyond the scope of the present study, it is felt that the issue should be looked into in some detail to understand the role it may play in assessing the effects of policy making support.

Two extreme positions can be discerned as far as the evaluation of a process is concerned: an idealistic and a realistic position.

According to the first, following David Hume's Treatise (1777), processes such as movement do not exist in reality and are just an imagination of the human mind. As a consequence, there is no point in trying to evaluate a process independently from a its products. Movement, for instance, only exists in the eye of the beholder, and should be evaluated by looking at individual consecutive measures of time and place Applying this position to the evaluation of the effects of participative policy modelling, the conclusion based on the present study's lack of outcomes would not be that a different methodological approach to the evaluation of the effects should be taken (there is no such thing as a policy making support process separated from the products it is based on). However, the conclusion would be that different dimensions to measure the products on should be included in order to be able to fully assess what is taking place during the sessions and how it may contribute to what policy makers do. So if no difference between pre- and posttest scores is found, despite the belief that some contribution has been made, it may well be that different dimensions should have been looked at. Dimensions such as 'the status of the knowledge participants have' (recall, it is possible that participants have changed without changing the content of their conceptualization for instance because they now feel they know for sure what the domain-specific elements to be included look like). In addition to such dimensions, it may also be that other, non-knowledge based dimensions have to be looked upon in order to understand how and to what extent participative policy modelling is contributing to the policy making process Note, that the bottom-line of this position is that, although it may well be that additional dimensions have to be focused upon in future studies, the evaluation of the effects can (in principle) be carried out by looking at product-scores on each of these dimensions. In the next section, discussing the implications the outcomes of the present study have from an

---

164 It is here that the difference between policy making and executive decision making becomes important for it is only in the former that the governing relationships are modified whereas in executive decision making one is primarily concerned with "*maintaining through time a complex pattern of relationships in accordance with standards or within limits which somehow come to be set as governing relations*" (Vickers, 1965, p 27). As a consequence, it is expected that if executive decision makers are invited out to take part in the policy modelling program rather than policy makers, one would not expect major changes in the definition and conceptualization of the problem for executive decision makers are not primarily concerned with (re)defining their situation by changing the assumptions underlying it (Klabbers, 1986)

165 Note however, that the evaluation carried out by Vennix has been product-oriented rather than process-oriented The very same can be said of the present study

'Actor approach' point of view, it will be argued that policy makers change the rules of the game they play while playing. The consequence of this would be that a product-based evaluation would have to be able to deal with all sorts of (un)expected changes in order to be able to assess what is taking place - all possible dimensions may have to be included on the pretest measure to be able to assess a change on that dimension. This would put a high burden on both the researcher (to consider all possible relevant dimensions and design a measurement instrument assessing all dimensions) and the participants, for such an instrument is likely to be like a Moloch. An intermediary position would be to include only those dimensions (including non-cognitive ones as well) that are felt to be the most important in the 'process' of policy making and policy making support assuming that one is able to determine in advance what course the actions/change will take during the participative policy modelling sessions.

According to the second position - a realistic one - there is something like a process independently (or differently) from the the products, the values there are at a particular point in time. Just like one can argue that movement is not the same as the distance divided by time (from an insider's point of view it is not the same, from an outsider's point of view it is), one can argue that in order to understand and appreciate the dynamic character of the processes that play a role during the sessions and in the policy making process, a different approach should be taken. Just like stopping the time will make 'movement' disappear, the process of interaction and constantly changing the meaning and use of the knowledge being arrived at in the sessions and the policy making process will disappear if only account is taken of the products that are, every now and then, produced by it. Hence, a completely different approach should be taken in order to understand and evaluate the impact participative policy modelling is having on the process of policy making.

Although the present study does not aim to solve this theoretical and methodological issue, section 8.4 will be concerned with the methodological implications of looking upon the policy making process as a process and taking into account the context-sensitivity of the knowledge that is arrived at in the sessions. However, as will be explained in more detail in that section, it will add to this a change in perspective. A change from an outsider's to an insider's perspective, in order to be able to understand the process of policy making support.

For now it suffice to conclude that on the basis of the lack of outcomes of the present study, the need for a different view on the *theoretical* aspects of evaluating participative policy modelling -'what' should be looked at - has become visible. In addition to this, and related to this change in what should be looked at, it has been clear that a critical reconsideration of the *methodology* - the 'how' - used for this should take place as well.

In the next session, the critical review of the 'what' of study's that concern the evaluation of the effects of participative policy modelling on the policy making process will be taken a step further by shifting from a Non-Trivial Machine approach to an Actor approach. Obviously, such a change also affects the way in which future evaluation studies will have to be carried out, that is, it affects their methodology.

### 8.3.3 IMPLICATIONS FROM AN 'ACTOR APPROACH' POINT OF VIEW

In the previous section, it was argued that in order to understand and explain the (lack) of effects found in this study, a change from a Trivial Machine approach to a Non-Trivial Machine approach should be taken. This to be able to take into account the situationality of the knowledge being constructed and used in the policy making process and the process of policy making support, and to consider the process-side of supporting the process of policy making.

By looking at the effects brought about from a Non-Trivial Machine point of view, we were able to detect the processes that are playing an important role in the context of the participative policy modelling program and which affect the meaning and use of the knowledge arrived at during the sessions. By stressing the process side of policy making, we have become aware that the outcomes should not be considered as products that can be delivered at a particular point in time and can be pre-defined, but as elements that play a role in and may affect the course of the activities that are taking place in the 'policy making situation' the participants are in[166].

Although the metaphor of the Non-Trivial Machine has made us realize that the knowledge used in the policy making process is context-sensitive, and should not be looked upon as a fixed product to be delivered, it does not question the participative policy modelling's aim to cognitively support the policy making process. In other words, it is suggested that one additional step may have to be taken to fully understand the effect of policy making support - to fully understand the present study's outcomes.

By looking more carefully at the policy modelling process taking place in the sessions, the question arises whether knowledge is the aspect of the policy making process that has to be focused upon. So far system dynamicists have been concerned with providing conceptual support, that is, providing knowledge that can be of use in the policy making process. From Watt's (1977) 'Why won't anyone believe us?', to Vennix's (1990) participative policy modelling program, attempts have been made to improve the utilization of the knowledge, by adjusting the program, decreasing the size of the models to deal with, and getting the clients involved in the process as much as possible. However, lack of success, that is, lack of changes found in the participants' conceptualization of the problem may have to be attributed to the fact that knowledge is not as important in the policy making process as was expected and that hence, irrespective of the quality or participativeness of the program, use of it will always be limited. In the early stages of the policy modelling process, social systems (groups of individuals) are not primarily concerned with knowledge and knowledge production, that is, with a rational approach to solving the problem, but with positioning oneself, with power and possible conflicts of interest, and only to a limited extent with knowledge: *"From our point of view policy formation is considered a continuous process of sorting out, arranging, maneuvering, wheeling and dealing of many actors who all are trying to cope with the ambiguities and equivocalities involved* (Klabbers, 1985, p. 146). Rather than being goal seeking, human systems are primarily concerned with building and maintaining relationships (Senge, 1992; Vickers, 1965). Or as Doerner (1983) puts it, policy making support should not be concerned with primarily providing cognitively-based support, but with steering, that is, giving structure and meaning to the processes that are taking place and the people that are involved in it: to develop the system that they are part of. However, because these elements of the policy making process *"do not fit into a positivist strategy and into the related tradition of "scientific analysis""*[167] (Klabbers, 1985, p. 146), they are commonly ignored in the design and evaluation of policy making support. It is only when a change in perspective is taken (from a Non-Trivial Machine to an Actor point of view) that the importance of those elements become visible and that consequently their influence on the policy making outcomes can be taken into account. Before that, such elements are often

166 The participative policy modelling method itself can be considered as a Non-Trivial Machine in that it enables participants to understand that changes over time (i.e. processes) are important in considering policy problems. Participants are taught that initial (historical) values affect the outcomes of the processes and that different initial values may result in the same effects. As such it will give those who have taken part in it the impression that opening the lid of the black box, i.e. examining the structure of the relationships that is responsible for the system's behaviour, will contribute to their understanding of the system and will enable them to predict the system's outcomes. The approach taken in the evaluation of the participative policy modelling method by contrast, has been more of a Trivial Machine nature. Recall that pre- and posttest scores were used to assess the effects brought about by the participative policy modelling sessions. Only limited account has been given of the contxt and processes that have affected the study's outcomes.

167 Or as Cowan, 1992, p. 396 puts it: "We may be restricting our perspective on the difference executives make, because of our predominant western-world paradigm (...). This paradigm is often characterized by short-term linear thinking, with expectations of immediate effects". If such a paradigm pervades research efforts, ...we may never attend to the processes that enable the linkages to occur".

regarded as interesting side-effects (e.g. Larsen, Morecroft & Murphy, 1991; Vennix, 1990) rather than belonging to the essence of policy making and hence policy making support.

To be able to perceive and appreciate the constructive or self-referential character of the policy making process and the processes that are taking place during the participative policy modelling sessions, it is suggested to change one's perspective from a Non-Trivial Machine approach to an Actor approach (Klabbers, 1990). This because the Actor approach allows one to take into account the social system's ability to change its own basic rules, to define (and maybe even construct) the problems within a particular context (people, time, and situation), whereas in the Non-Trivial Machine approach it is assumed that the basic rules remain relatively unchanged as a result of which it is, theoretically, possible to determine (and predict) the course of action taken by the system if one knows the systems initial values and basic rules.

In policy making in contrast, the basic rules themselves can and constantly are changed in the process of defining the system and defining and legitimating oneself in it. To illustrate this process, recall that our participants did not simply accept the program offered to them, but tried to change it according to their own interests. An experience shared by for instance Lane (1992a), who found himself in a position in which participants decided to select and elaborate on an option that had not been prepared by the expert modellers.

Hence, to account for the absence of any major differences between pre- and posttest, it is suggested to go one step further than the Non-Trivial Machine approach. It is claimed that because policy making is not primarily concerned with knowledge and cognition per se, but also with the meaning, applicability, and situationality of this knowledge, participative policy modelling should not be primarily concerned with providing knowledge-based support. This obviously affects the evaluation of the participative policy modelling method in that one may have to focus on other aspects of the process than knowledge alone. Rather than considering participative policy modelling as a learning process, it is suggested to look upon it as a self-referential or self-constructive process for it allows participants to define and construct their own reality and their own place in it. Policy makers do not react to stimuli in their environment, recognized as problems, but define and create the problems themselves in interaction with those stimuli and other actors. Or as Stone (1987, p. 116) puts it:

> "Problems are not "given, out there in the world waiting for smart analysts to come and define them correctly" but "created in the minds of the citizens by other citizens, leaders, organizations, and government agencies, as an essential part of political maneuvering."

Or as Felling (1974) puts it, to understand the process of policy making, account should be given to the communcation and information networks that affect and are affected by the policy making process. Policy-making is not taking place in a social vacuum.

However, this is not to say that problems (or situations) are the product of policy makers only, that they can define and construct the system that they are part of completely, as Weick (1979, p. 44) puts it citing Baldwin:

> "The will of a general may stimulate his troops and so bring to him the victory he believes in; but such an act of the general's will cannot replenish the short supply of powder or shells, on which the issue of the battle perhaps more fundamentally depends." (Baldwin, 1909, pp 72-73)

Despite this, it is believed that within particular boundaries, policy makers are concerned with the construction of their reality, and the definition of their role in this, and that consequently participative policy modelling can be considered as a means to assist them in doing so, for it enables policy makers to determine the people (actors), rules and dynamics of the situation they are dealing with, and help them in the process of re-constructing this reality.

Because of this constructive process in which the basic rules of the reality one feels one is in, can be (and often are) changed as part of the process, it is questionable as to whether we can determine what the outcomes of the intervention are, and to what extent they can be attributed to our intervention. Or as Von Foerster (1984, p. 12) puts it: *"There exists a large class of machines whose driving and state functions are such that it is in principle impossible to infer these functions from the results of a finit number of tests...This also means that there are non-trivial machines that are unknowable"*, In other words, without looking into the black box (knowing the system's internal state, its history or initial values) one cannot infer the relationship that exists between in- and output. Let alone when the relationships within the system are not stable and change over time as well as is the case with social systems. It is almost like raising children: you give them input and simply hope for the best - hope that they will use the input in the process of becoming the ones that they are, will use it in the self-referential process of becoming the individuals they are.

However, not only does such input not guarantee any specific outcomes, it also is extremely difficult to tell where and how the input was used in the self-referential process. The very same conclusion seems to hold for the participative policy modelling intervention in the context of policy making support. It does seem to contribute to the process of re-defining and re-constructing oneself as an organization and as individuals (participants say it has been useful), but it is difficult to determine where exactly, how, and to what extent. In the next section, some of the methodological implications of this conclusion will be discussed in more detail, for it is clear that changing the role of participative policy modelling in the policy making process, does affect the way in which these effects should be evaluated. Moreover, some recommendations will be made with respect to the participative policy modelling method itself, for it may well be that changes can be made to the program to improve its effectiveness in bringing about a self-referential process[168] rather than the cognitive changes it once was designed for.

To understand the developments that have taken place in the relationship between participative policy modelling and policy making, a brief summary will be given of the three stages that have been discerned. The differences between the three metaphors (Trivial Machine, Non-Trivial Machine, and Actor approach) and their consequences for the way in which one can look upon participative policy modelling and the contribution participative policy modelling can make to the policy modelling process, can be depicted as follows:

[168] Supporting the policy making process by having policy makers take part in participative policy modelling sessions can be considered as self-referential processes in that the knowledge of the system under consideration is affecting the system under considerationitself. This because the policy makers taking part in the program are part of the system as well. Hence by referring to an external object, they actually refer to a system of which they are part of themselves. Moreover, their (increased) understanding of the structure of the model (and the system itself) also affects the structure of the model (and system) itself. They actually change, (re)construct the world that they act in. Hence, the distinction between 'virtual' or constructed world (the model under consideration) and the word that they actually live and work in, is gradually disappearing (Klabbers, 1990).

'Actor approach' point of view, it will be argued that policy makers change the rules of the game they play while playing. The consequence of this would be that a product-based evaluation would have to be able to deal with all sorts of (un)expected changes in order to be able to assess what is taking place - all possible dimensions may have to be included on the pretest measure to be able to assess a change on that dimension. This would put a high burden on both the researcher (to consider all possible relevant dimensions and design a measurement instrument assessing all dimensions) and the participants, for such an instrument is likely to be like a Moloch. An intermediary position would be to include only those dimensions (including non-cognitive ones as well) that are felt to be the most important in the 'process' of policy making and policy making support assuming that one is able to determine in advance what course the actions/change will take during the participative policy modelling sessions.

According to the second position - a realistic one - there is something like a process independently (or differently) from the the products, the values there are at a particular point in time. Just like one can argue that movement is not the same as the distance divided by time (from an insider's point of view it is not the same, from an outsider's point of view it is), one can argue that in order to understand and appreciate the dynamic character of the processes that play a role during the sessions and in the policy making process, a different approach should be taken. Just like stopping the time will make 'movement' disappear, the process of interaction and constantly changing the meaning and use of the knowledge being arrived at in the sessions and the policy making process will disappear if only account is taken of the products that are, every now and then, produced by it. Hence, a completely different approach should be taken in order to understand and evaluate the impact participative policy modelling is having on the process of policy making.

Although the present study does not aim to solve this theoretical and methodological issue, section 8.4 will be concerned with the methodological implications of looking upon the policy making process as a process and taking into account the context-sensitivity of the knowledge that is arrived at in the sessions. However, as will be explained in more detail in that section, it will add to this a change in perspective. A change from an outsider's to an insider's perspective, in order to be able to understand the process of policy making support.

For now it suffice to conclude that on the basis of the lack of outcomes of the present study, the need for a different view on the *theoretical* aspects of evaluating participative policy modelling -'what' should be looked at - has become visible. In addition to this, and related to this change in what should be looked at, it has been clear that a critical reconsideration of the *methodology* - the 'how' - used for this should take place as well.

In the next session, the critical review of the 'what' of study's that concern the evaluation of the effects of participative policy modelling on the policy making process will be taken a step further by shifting from a Non-Trivial Machine approach to an Actor approach. Obviously, such a change also affects the way in which future evaluation studies will have to be carried out, that is, it affects their methodology.

### 8.3.3 IMPLICATIONS FROM AN 'ACTOR APPROACH' POINT OF VIEW

In the previous section, it was argued that in order to understand and explain the (lack) of effects found in this study, a change from a Trivial Machine approach to a Non-Trivial Machine approach should be taken. This to be able to take into account the situationality of the knowledge being constructed and used in the policy making process and the process of policy making support, and to consider the process-side of supporting the process of policy making.

By looking at the effects brought about from a Non-Trivial Machine point of view, we were able to detect the processes that are playing an important role in the context of the participative policy modelling program and which affect the meaning and use of the as well[170]: policy making considered from an 'Actor approach' point of view is looked upon as a kind of self-steering, a recursive processing of information, i.e. conversation and communication through which those involved in it (policy makers) continuously transform their collective structure (Klabbers, 1990). This now, is the reason why participative policy modelling sessions should be carried out in a participative way. Not because of the knowledge that can be arrived at, but because the construction of the social system they are part of is taking place in these sessions. Without their presence, no such construction can take place. The kind of learning environment in which such (re-) construction of social reality can be carried out, is called a self-organizing learning environment (Klabbers, 1990, p. 10).

Summarizing the above discussion, by changing one's perspective from a Non-Trivial Machine to an Actor point of view, the (lack of) outcomes of the present study can be interpreted differently. By adopting an Actor point of view, we have become in the position to not trivialize the relationship between treatment offered and outcomes expected. It is recognized in the Actor approach that the basic rules of the social system that is taking part in the program can (and probably will) change during the sessions: participants constantly change, construct, and legitimate the system they are part of. Hence, to assess the contribution of participative policy modelling to the policy making process, it is believed that the knowledge-based approach itself may have to be questioned because no major cognitive changes were found in the present study, and we have come to realize the restrictedness of a rationalistic definition of the policy making process according to which knowledge is considered to be one of the most important elements and the policy making process is being defined as goal-oriented in nature.

How these changes in the way in which participative policy modelling is supposed to contribute to the policy making process affect the methodological issues related to the evaluation of its effects and the way in which future participative policy modelling programs should (or could) be designed, will be discussed in the sections that follow.

## 8.4 A METHODOLOGICAL REVIEW: THE 'HOW' OF POLICY MAKING SUPPORT EVALUATION

Having discussed the way in which the 'what' of evaluative studies regarding the effects of participative policy modelling may have to be changed to do right to both the method and the policy making process it aims to contribute to, account should be taken of the methodological implications these changes may have for evaluative studies to be. This because changing the contribution participative policy modelling may have on the policy making process, will almost inevitably change the evaluation criteria used to measure such a contribution and the approach taken to do so.

As described in the previous section, it seems important to take both the **self-referential** character of the policy making process and the processes that are taking place during the participative policy modelling sessions (i.e., the process of reconstructing the system and thus reconstructing oneself, for the distinction between the system under consideration and actor observing the system is vanishing) <u>and</u> the **process** character of policy making itself into account when considering the process from an Actor point of view. The question then becomes, how to evaluate a process to start with, and more importantly, how to evaluate a process that is self-referential in nature.

In the present section, some methodological issues will be touched upon. Note however, that a full-fledge description of how to measure these contributions will not be given for such an examination falls beyond the scope of the present study. Moreover, it is expected to be probably more difficult to come up with methodological guidelines on how to assess such a contribution than it will be to determine exactly where the contribution of participative policy modelling to the policy modelling process would be

170 As Morgan (1986, p. 236) puts it: "Thus a system's interaction with "its environment" is really a reflection to part of its own organization".

from a theoretical point of view[171]. It is expected that careful analysis will result in a list of criteria that can be used to evaluate the contribution participative policy modelling can have on the policy making process. However, meeting the methodological requirements needed to measure this contribution may be more of a problem. It seems that we may have to come up with new methodologies for that: ".... *has occurred at various times in the physical sciences, whereby once investigators realized that something may exist that may be worth exploring, they developed ways to find out, which led to discovery and advancements in knowledge*" (Cowan, 1992, p. 399).

Prior to discussing some of the methodological implications that follow the changes in the 'what' of studies to be described in the previous section, a brief summary of the most important features of this change will be given first.

On a high level of abstraction, these changes in the 'what' of future research concerning the effects of participative policy modelling are the following:

* Focus on the processes that are taking place during policy making support inerventions (and products) rather than the products only
* Assess the degree to which participative policy modelling has been able to support the self-referential processes taking place in organizations and policy making processes rather than on cognitively changes predominantly.
* Take into account the context in which these processes are taking place. This to be in the position to understand how changes in the context (may) have affected the (re-) construction of the system during the sessions.

### *Methodological implications*

One of the most important methodological implications of the above described changes in the 'what' of evaluations to be, is that future studies should take account of the so-called self-referential processes rather than knowledge or conceptualization in isolation. The question then becomes, how can and should such self-referential processes be assessed? Again, although a full-fledged answer to this question will not be given because it falls beyond the scope of the present study (making us realize that such processes are important and may have to be included in future studies evaluating the effects of participative policy modelling falls *within* the present study's scope), some ideas will be presented that may serve as an input for future studies.

Firstly, to perceive and understand the processes that are taking place, it is clear that one cannot remain an outsider and observe the processes from an observer's point of view, for instance by comparing in- and output scores. This because from an outsider's point of view, one would not be able to see and understand the self-referential character of the processes, the fact that changes in the knowledge and meaning of the system is also affecting the system itself. One would not see the use (and misuse?) of the program offered and the changes in the rules made by the participants[172]. It is only when one becomes an insider as well, that one is able to get an idea of what is going on during the sessions and how the program has been of any help to the organization or the individuals taking part in it. It is only when one has knowledge of the inside of the system, that one is

171 This is not to say that no further examination of the criteria by means of which the contribution of participative policy modelling to the policy making process is required. By contrast, it is felt that one of the most important things to examine in detail is the relationship between participative policy modelling and policy making from an Actor point of view. It is only when one has a clear idea of the theoretical issues (the 'what') involved in such a study, that one can consider methodological issues, issues related to the 'how' of such an evaluation.

172 One of the important contributions of the participative policy modelling method may well be that it provides the participants the opportunity to change the rules of the game, that is, the rules of their reality in a pre-structured and time-compressed way, thereby speeding up the process of changing one's organization in light of things to be.

able to see the dynamics of the system on a micro-level; that one is able to have knowledge of the system in its situational context - the people attending, the situation those people are in, and the time at which the program is taking place.

One of the methodological implications of such a change in role as researcher obviously is that the distinction between object of research and researcher itself will vanish. The researcher no longer is an actor who is independent of the object of his or her research, the researcher's presence no longer is neutral, and (s)he is also affecting the course of actions taken in the program itself[173] - researchers become an element of the system they have under consideration, thereby affecting the system one is 'observing'. Or as Cowan (1992, p. 398) puts it:

> "Though some of the gap in studying real-time thinking may be attributed to the short-term paradigm mentioned previously, some may be attributed to difficulties inherent in measuring it. Particular problematic is the fact that real-time thinking may be altered when studied with intrusive measures or when the context is altered."

Roles that a researcher can take in this process are for instance 'designer of the program', 'interpreter of the processes and outcomes', 'negotiator', and 'facilitator' (Klabbers & Scheper, 1988).

The implication of this clearly is that the difference between external consultant and client is disappearing. Consultants / researchers are expected to become insider's as well to be able to understand the role they have and the contribution they make to the organization and to perceive and understand the processes that are taking place. In addition to this, it is important that clients also change their perspective, and every now and then become an observer, and outsider rather than stay an insider at all times. It is in this change of perspective (from outsider to insider and vice versa) that one starts to understand the processes that are taking place and is able to consciously change the course of actions.

> "An outside actor can consider it [the social system] an object. An actor may also choose to take a boundary position, that is, a balance between a mode of participation (inside) and reflection (outside). This changing of perspective somehow resembles boarding a helicopter. When it takes off, it rises above the level of individual actors who are engaged in numerous double-inter-acts. We will lose detail at the micro-level but gain perspective at the level of patterns of double-inter-acts. We become aware of the structure and boundaries of the collective structure that gives shape to the social system. After landing, we are again inside actors involved in the various double-inter-acts". (Klabbers, 1986, pp. 66-67)

One of the advantages of having client participate in the construction of the model of their reality (and thereby constructing their reality as well) is that is puts them into the position of both an **insider** (they get involved and occasionally carried away in the process of constructing their own reality) and an **outsider**, a researcher, who is responsible for the pre-structured process. Having clients participate in the design of the program and be responsible for that part of the process as well, forces them to take both positions either at the same time or at different points in time.

The second methodological implication that should be taken seriously is based on the observation that policy making processes have no closure and that the processes that are taking place are more important than their specific outcomes at a particular point of time.

---

173 Zouwen, van der (1988, pp. 86-87) for instance, describes the way in which researchers involved in a survey adjust the surveys depending on factors such as the quality of the interviewer and the quality of the information provided in the interviews. As anticipating researchers rather than neutral observers, they try to steer the processes and hence change from a traditional observer's point of view to an actor or participant's point of view.

This because it is difficult to tell from an outsider's point of view, trying to contain the process and stop it at a particular time to measure its value, whether anything important has changed or what eventually the contribution of the program offered to the policy making process will be. To illustrate the difficulties of a product or data-model approach, it is for instance possible that participants reconsider their position and decide not to change (they may feel more confident as to what to do in the situation). Such a change would not become visible if attention is paid only to products or data-models rather than the processes that have taken place. One of the steps that may have to be taken in order to be able to assess the impact participative policy modelling is having on the self-referential processes that are taking place in an organization, is to make use of so-called process-models, in which, based on the structure of the situation people are in, understanding can be acquired about the processes that are taking place (Berting, 1988; Buffart, 1988; Klabbers, 1988).

Hence, to assess the contribution participative policy modelling has made to the self-referential processes, one should not use the classical pre- and posttest designs. Not only because one can question the use of data-models to arrive at an understanding of the processes, but also because self-reference is related to feedback of the dependent variable to the independent variable, which cannot be made visible in an experimental design (Zouwen, van der, 1988). In other words, because the treatment (independent variable) offered to the participants (the participative policy modelling method) not only affects the participants taking part in it (the dependent variable) but itself is also affected by the people taking part in it (the direction of causality is from dependent to independent variable as well), one cannot make the recursive processes taking place between the program offered and the system taking part in it visible by means of an experimental pre- and posttest design. It is suggested by for instance, von Foerster (1984) and Klabbers (1988), that a different kind of concept should be used to describe and understand these recursive processes. Rather than what they call the traditional first-order concepts (i.e. concepts that refer to 'external' objects / realities), use should be made of second-order concepts. These second-order concepts seem promising to describe the self-referential nature of the processes that are taking place in the sessions because they not only refer to the external object or external reality, but also to the actor using the concept him- or herself.

> Let me begin with "purpose". If taken as a first-order concept one may speak of something "having a purpose". However, taken on its second-order level we may ask "what is the purpose of 'purpose'?", that is, to ask why introduce the notion of purpose in the first place...However, by paying attention to the autological nature of "purpose", our gaze is shifted from "something" , the observed, to "somebody", i.e., the one who uses the term, that is, the observer. (von Foerster, 1984, p. 4)

By using second-order concepts, one is able to bridge the gap that normally exist between researcher/observer and object of research, thus doing right to the self-referential character of the processes under consideration and the double nature of examining such processes (being both an outsider and an insider at the same time or at various times). Note that the self-referential nature also affects the role the researcher is playing. (S)he no longer is remaining a neutral observer or outsider, but affects and is affected by the system under consideration.

In addition to the awareness that a different kind of design than an experimental pre- and posttest design is needed to be able to make self-referential processes visible, one could also try to come up with process variables rather than product variables to acquire some understanding of the processes that have taken place during the sessions. Examples of such process-indicators may for instance be 'the number of questions asked', and 'the number of interactions taken place between particular actors'. However, bear in mind that it still is not possible to know what the meaning of these process variables is, if not an insider's perspective is taken. Hence, the most important recommendations at this stage seem the change in perspective that both researchers and participants should take in order to understand what has taken place in the sessions and the need for other approaches to evaluation than the 'classical' product- or objective-orientated evaluations:

> "Undoubtedly, the open-endedness which is the very heart of many simulation games presents a special problem for an objective-orientated evaluation since goals may be tentative, speculative and probabilistic." (Barnett, 1984, p. 169)

Hence, alternatives to the classical paradigm are needed to illuminate the evaluation of this complex and complicated method of policy making support. However, as Stenhouse (1975) puts it, it is necessary to be aware of the danger of superficiality inherent in a completely anthropological approach in which the processes that have taken place are described and interpreted. *"Therefore there is a need to continue along both the interpretive and objective-orientated routes"* (Barnett, 1984, p. 171). It is because of this concern for a purely 'anthropological' or 'soft' approach that the introduction of a second-order formal system by von Foerster may be a promising step forward to understanding .

Based on the above discussion, it is felt that future studies should not only be concerned with a careful examination of the contribution participative policy modelling is expected to make to the policy modelling process (the 'what' of evaluations to be), but also determine how such studies should be carried out in order to be able measure this contribution. As such, we follow Schein (1993, p. 85) who states: *"I am stuck by how little we really know about the dynamics of organizations and social systems, and how little we know about the learning process"*, but would like to add: "We are stuck by how little we really know about how to *evaluate* the dynamics of organizations and social systems, and how little we know about *evaluating* the learning process".

Having discussed the implications of the present study's findings for both the 'what' and 'how' of future evaluations, an overview will be given of the way in which it is believed the participative policy modelling method itself can be improved in light of these implications. Changing one's perspective from a Non-Trivial machine approach to an Actor approach will clearly affect the way in which participative policy modelling sessions are to be designed.

## 8.5 IMPLICATIONS FOR THE DESIGN OF PARTICIPATIVE POLICY MODELLING

In the present section, the recommendations that are based on the experiences acquired in the present study within the framework of *conceptually* or cognitively supporting the policy making process will be presented first. These recommendations will be followed by a critical review of the participative policy modelling program in light of the change in perspective discussed in the above sections.

### *Improving the design of participative policy modelling programs from a 'traditional' point of view*

One of the important recommendations made in the Vennix (1990) study concerned the degree to which participants are able to challenge the model created by expert modellers. Following Vennix, it was decided to provide participants in the present study with the opportunity to do so. However, as explained above, hardly any changes were made to the so-called preliminary model. It was our impression that the relatively low number of changes should not necessarily be considered as an indication of the high quality of the model, but could well be attributed to a lack of involvement on the participants side. Hence, it is our recommendation that future version of a participative policy modelling program are participative in a true sense in that the model is created by the participants themselves rather than by one or more expert modellers. Such high level of participativeness has for instance already been accomplished in a project reported by Larsen, Morecroft, and Murphy (1991).

Obviously such a change in the program requires participants to invest more time in the program. It is our experience that one simply cannot expect clients (with a very divert background) to be able to acquire a thorough understanding of the system they are part of in say three of four sessions three-hour sessions.

In addition to this, it is important that more time is spent on either specific domain-specific or strategic issues rather than covering both aspects of conceptualizing a particular policy problem at the same time. Looking back at the design of the participative policy modelling method, it is our impression that more time could and should have been devoted to the design of the program in relationship to the learning objectives we had in mind.

As far as the use of the computer is concerned, the present study followed Vennix by limiting the use of the computer as much as possible. Not only to avoid the so-called 'video-game phenomenon' (Senge, 1989, p. 233), but also because using a computer may put some people off (some participants are unable to, not willing to, or afraid to work with a computer), or may lead to too much confidence in the model (some people seem to be more likely to accept the model when it is presented as a computer model). In addition to this, we found that policy makers enjoy talking and argueing about the model much more than doing computer analyses. One reason for this may be that the participants did not have the command of the model required to enjoy and appreciate the analyses. Another reason is that they may have considered communication the most important objective for taking part in the session, as a result of which they looked upon the computer runs as of secondary importance. It may also be because, as argued above, policy making and policy making support is much more related to maintaining and building relationships than with acquiring knowledge about the structure and dynamics of a particular computer model. Irrespective of the reasons why people enjoyed discussing the conceptual model much more than using the computer model, the fact that they did should be taken into account when designing future participative policy modelling programs.

Regarding the flexibility of the program, it is important to note that since use was made of the DYNAMO software package, only limited changes to both the model and the analyses to be carried out could be made. Future applications should definitely try to overcome this limitation in order to be able to deal with client's wishes more flexible (e.g. what analyses to carry out, what variables to compare and how to adjust the model).

Although not really an aspect of the design of a participative policy modelling program, one should make sure that clients have a clear idea of the objectives of the program and at least a rough idea of what the program is going to be. In the present study, an introductory meeting was held with all the people who would participate in the program, to explain the program's background, objectives, and content. Despite this meeting it turned out that particularly with respect to the program's objectives, different point of views existed among those taking part in the program. As discussed above, these differences in perspective have affected the participants perception and appreciation of the program, and probably their willingness to invest time in it as well.

Regarding the preparatory text given to the participant to get introduced to the preliminary model, the participant-based evaluations together with comments made by the participants during the sessions, clearly indicate that preparatory texts should be very concise, if given at all. In light of the degree of participation discussed above, it is recommended that no preliminary model is created and that participants are given the opportunity to define their own problem and subsequently construct their own 'preliminary' model.

In the present study it was felt that the construction and discussion of the model was of a rather non-committal, open-ended nature for no action plan had to result from it. Since such an academic character may result in lack of commitment and involvement, it is suggested that future studies should be more related to the organization's policy plan or should at least come up with a kind of action plan or steps to be taken.

### * *Improving the design of participative policy modelling programs from a NTM and/or Actor point of view*

Since a participative policy modelling program from a Non-Trivial Machine and/or Actor point of view should be concerned with self-referential processes (i.e. re-construct their own reality, a reality of which they are part) rather than the conceptualization of an 'external' problem only, participative policy modelling programs may have to be designed differently. Since a detailed description of what exactly the contribution of participative policy modelling to these processes can be, has not been given in the present study, the recommendations for future participative policy modelling programs will be of a general level as well.

To start with, it is important that participative policy modelling programs allow participants to change their perspective, either by moving from an insider's point of view to an outsider's point of view and vice versa, or by playing a role that is not necessarily the one participants plays in reality. As such, participative policy modelling should include elements from the world of simulation/gaming. Note that this recommendation is in line with the study carried out by Vennix (1990), for in the Vennix study students played roles such as the Minister and Secretary of the Department of Social Security (Vennix, 1990, p. 88).

Another way to force participants change their perspective, is by making this an element of the debriefing section. By asking the participants to step aside the construction of the model and the processes they were involved in while constructing it, one allows them to move from an insider's to an outsider's point of view.

Moreover, participants should be enabled to construct their own reality rather than forcing them to comply with all sort of rules (e.g. steps they have to play in the sessions in a particular order). As such participative policy modelling should incorporate elements of the so-called free form or frame games in which participants are given very little instructions and rules to start with and construct their own rules during the sessions (Klabbers, 1990).

As far as the participation is concerned, it should be clear by now that maximum participation should be strived for, because of the constructive processes that are taking place in the construction of the model. Consequently, preparatory activities should be removed from the program. However, as far as the use of a computer is concerned, it is not our believe at this stage that it should be limited or abandoned from the program altogether for computer analyses can be of great help examining the dynamics of the model/reality constructed in the program.

Finally, regarding the way in which the program should be designed to enable researchers to change perspective as well, one could for instance make use of two rather than one facilitators. If two facilitators attend the sessions, one for instance, can take part in the session as a participant (e.g. by giving that person a coordinating role in the session, or having him/her carry out computer manipulations when needed) so that the other facilitator can take on a more distant role - observing the processes that are taking place rather than taking part in it. Alternatively, one could have a client participant working closely with the expert modellers and taking part in the sessions, who could take a boundary position in between an insider's and an outsider's point of view, thereby assisting the expert modeller in the process of understanding and debriefing the processes that are taking place.

## 8.6 PRACTICAL IMPLICATIONS

This study has shown that the evaluation of participative policy modelling still has a long way to go.

To start with, the discussion of the outcomes and the introduction of an alternative theoretical framework to account for these effects has provided a somewhat different perspective on the relationship between participative policy modelling and the policy making process. The critical reflection carried out in the present chapter has made us realize that other aspects than knowledge and understanding of the policy making problem at hand determine the course of action in the policy making process - aspects that concern the (re-)construction of the reality of those taking part in it.

As a consequence, future studies will have to examine the relationship between participative policy modelling and the policy making process in detail first, before they will be in the position to determine how such research will have to be carried out.

However, as stated before, it is expected that the construction of a methodologically appropriate approach to assess the impact participative policy modelling is having on the policy making process will be even more difficult than deciding where exactly the contribution of the method to the policy making process will be.

Hence, it is felt that now that we have seen that the cognitive framework based on a rationalistic approach is probably not sufficient to account for the contribution participative policy modelling is having on the policy making process, further empirical research should be postponed. Rather than carrying out another empirically-based evaluation of the effects participative policy modelling is having, additional theoretical research should be done first. It is only when the theoretical framework has been (re-) constructed, and an appropriate methodological approach has been developed, that new empirical studies should be carried out.

# CHAPTER 9: SUMMARY

In this chapter, a brief overview will be given of the study's most important findings and the conclusions that have been arrived at on the basis of these findings. It will start with a brief summary of the approach taken to assess the effects of participative policy modelling on the conceptualization of a problem. In addition to this, it will briefly describe the organization and individual's taking part in the program. This will be followed by the presentation of the five research questions that have been focused upon, and the major outcomes with respect to each of these research questions. Finally, it will summarize the main conclusions that have been drawn on the basis of these outcomes. Included in this summary will also be the critical reflection of the approasch taken in the present study itself and the introduction of an alternative theoretical framework in the previous chapter.

## Evaluating the effects of participative policy modelling

To assess the effects of participative policy modelling, it was decided to use a one-group pretest-posttest design. Purpose of this design was to assess the effects participative policy modelling has on the way in which participants look upon or conceptualize a particular policy problem. As a consequence, participants were asked to answer three questions related to the costs of health care at the pre- and posttest. In addition to this, they were asked to fill out a questionnaire at the end of each session to express how they felt about the program.

## Participative policy modelling

The participative policy modelling program consisted of four sessions in which participants build, examine, and play with a system dynamics model of a particular problem. Each of the sessions took about three to four hours. Prior to attending a session, participants were asked to read a so-called workbook to prepare for it. Note that to speed up the building program, use was made of a preliminary conceptual model, that is, a model constructed by expert modellers.

## Participants

The organization taking part in the sessions was a regional health care insurance company. Twenty-two employees in total from both the medical and economical-legal departments were invited out to take part in the program. Of these twenty-two participants, eightteen participated to such an extent that they were able to qualify for the on-treatment group. The organization was willing to take part in the program for two reasons. Firstly to prepare for the changes they were expecting in the world of health care and health care insurance and secondly, because they wanted the two departments to be able to work more closely together, to get to know each other.

## Research questions

To assess the effects of participative policy modelling on the conceptualization of a policy problem, the present study examined as to whether participative policy modelling was able to increase:

- the *strategic* knowledge *individual* participants have of the problem at hand.
- the *domain-specific* knowledge *individual* participants have of the problem at hand.

- the *inter*-individual *strategic* knowledge participants as a group have of the problem at hand.

- the *inter*-individual *domain-specific* knowledge participants as a group have of the problem at hand.

- the *awareness* of the other participants' point of view

## Findings

With respect to all five research questions (with the exception of the fifth), no major changes were found. Despite the fact that our participants changed about 50 per cent of their conceptualization, no significant increase in the individual domain-specific and strategic knowledge was found. In addition to that, no increase in shared understanding (increase in knowledge and decrease in variance among the participants) could be concluded to. In contrast to this, the questionnaire-based findings did suggest some increase in knowledge of the problem, and clearly were in support of the claim that the program had been useful and interesting.

## Conclusions

To account for these somewhat surprising results, attention was paid to a number of elements where it was felt the present study could be improved. The use of one and the same instrument to evaluate a change in domain-specific and strategic knowledge, the lack of motivation to answer the same three questions two times, the limited attention that was paid to aspects such as time and delays, the use of a preliminary model, the lack of knowledge of how cognitive maps can change over time, the somewhat limited sophistication of the instrument used to assess changes in conceptualization, and the limited support that was given to the program by the CEO, may all have affected the outcomes negatively and thus be held responsible for the (lack) of effects concluded to in the study.

In addition to questioning the way in which the present study's theoretical framework has been employed, one can also question the framework itself and suggest an alternative point of view to examine the relationship between participative policy modelling and policy making support. Despite the fact that the above-mentioned limitations of the present study's approach suggest that there is room for improvement within the current approach, it was felt that the introduction of an alternative theoretical framework should not be postponed until a more 'ideal' evaluation program had been carried out. This because it may well be that these so-called limitations are in fact unavoidable characteristics or elements of the (social) systems we are dealing with and aim to evaluate.

The alternative theoretical framework that was introduced to account for the lack of change claims that one should not only focus on datamodels in the attempt to evaluate the effects of participative policy modelling, but also take into account the processes that have taken place. This because the knowledge that is constructed or acquired in the model-building process is of a situational nature - understanding the structure of the system under consideration may affect the system itself and the developments within the system (its historical value) affect the status of the knowledge one has of the system as well. As such, one has to take into account the self-referential nature of the processes that are taking place in understanding the system and the process of understanding the process of understanding the system. For this it is suggested that a change in perspective is needed; a change from oberserver or outsider to participant or insider.

It is important to realize that although the above outlined alternative theoretical framework seems to be promising to account for the effects found in this study, further research needs to be done regarding both the *theoretical* aspects (i.e. the 'what' of research to be) <u>and</u> *methodological* aspects (i.e. the 'how' of future research) of evaluating the contribution of participative policy modelling to the policy making processs.

# REFERENCES


Ackoff, R.A. (1969). Fundamentals of operations research. New York.

Ackoff, R.A. (1974). Redesigning the future: a systems approach to societal problems. New York.

Akkermans, H. (1992). Participative Modelling To Support Strategic Decision Making in Operations - A Case Study. System Dynamics Conference, Utrecht, 21-31.

Allen, T.H. (1978). New Methods in Social Science Research. Policy Science and Futures Research. New York: Praeger Publishers.

Anderson, C.R. (1970). An experiment on behavioral learning in a consumer credit game. Simulation and Gaming, 1, 1

Axelrod, R. (1976). Structure of decision: the cognitive map of elites. Princeton (NJ): Princeton University Press.

Bakken, B. (1989). Learning in Dynamic Simulation Games: Using Performance as a Measure In: P.M. Milling, and E.O.K. Zahn, [Eds]., Computer-Based Management of Complex Systems. Proceedings of the 1989 International Conference of the System Dynamics Society, Berlin: Springer, 309-316.

Barnett, T. (1984). Evaluations of simulations and games: a clarification. Simulation and Games for Learning, 14, No. 4, 164-175.

Basuray, M.T. (1982). The learning style inventory debate revisited: an empirical assessment of the construct validity issue related to experiential learning theory. In: Developments in business simulation and experiential exercises. Proceedings of the 9th annual conference of the Association for Business Simulation and Experiential Learning, febr., Phoenix, Arizona, 55-63.

Bereidheid tot verandering, advies van de Commissie Structuur en Financiering Gezondheidszorg [Willingmess to Change. Advice from the Structure and Finance of the Health Care System Committee]. Den Haag: Staatsuitgeverij., 1987.

Berting, J. (1988). Maatschappelijke veranderingen, theorievorming en onderzoeks-methoden in de sociale wetenschappen. Chapter 3 In: Klabbers, J.H.G. [Ed], Kennen en organiseren van informatie. Over machines en actoren. [Knowledge and the organization of information. About machines and actors]. Deventer: Van Loghum Slaterus.

Bex, J.H.M. (1989a). Structuur van de Nederlandse gezondheidszorg. Chapter 2 in: Reader: Organisatie gezondheidszorg en juridische aspecten. [Reader: Organization Health Care System and Legal Aspects]. Nijmegen.

Bex, J.H.M. (1989b). Kosten van de gezondheidszorg. Chapter 8 in: Reader: Organisatie gezondheidszorg en juridische aspecten. [Reader: Organization Health Care System and Legal Aspects]. Nijmegen.

Bonham, C.M., and Shapiro, M. (1976). Explanation of the Unexpected: The Syrian Intervention in Jordan in 1970. Chapter 6 in: Axelrod, R. (1976). Structure of decision: the cognitive map of elites. Princeton (NJ): Princeton University Press.

Bonham, C.M., and Shapiro, M., and Heradstveit, D. (1988). Group Cognition: using an oil policy game to validate a computer simulation. Simulation and Games, 19, No. 4, 379-407.

Boocock, S.S., and Schild, E.O. (1968). Introduction. In: Boocock, SS and Schild, E.O [Eds.], Simulation Games in Learning. Beverly Hills, CA: Safe.

Boot, J.M., and Knapen, M.H.J.M. (1984). De Nederlandse gezondheidszorg [The Dutch Health Care System]. Utrecht/Antwerpen: Het Spectrum.

Bredemeyer, M.E., and Greenblatt, C.S. (1981). The educational effectiveness of simulation games. A synthesis of findings. Simulation and Games, 12, No. 3, 307-332.

Breuer, K. Tennyson, R.D., and Lippert, R. (1989). Lehren und Lernen von kognitiven Strategien mit Problemorientierten Computersimulationen [Becoming knowledgeable and skilled in cognitive strategies by problem-oriented computer-simulation]. In: H. Levenig and W. Schöler [Eds], Kommunikation und Begegnung Reflexionen und pragmatische Ansätze der Pädagogik - Festschrift zum 75. Geburtstag von Johannes Zielinski. Heidelberg: Sauer-Verlag.

Breuer, K., and Kummer, R. (1990). Cognitive Effects from Process Learning with Computer-Based Simulations. Computers in Human Behavior, 6, 69-81.

Brewer, G.D. (1973). Politicians, bureaucrats and the consultant: a critique of military problem solving. Cambridge, MA: Harvard University Press.

Brewer, G.D., and deLeon, P. (1983). The foundations of policy analysis. Homewood, Illinois: The Dorsey Press.

Brown, J.S., Collins, A., and Duguid, P. (1989). Situated Cognition and the Culture of Learning. Educational Researcher, 18, No. 1, 32-42.

Buffart, H.F.J.M. (1988). Over de structuur van de niet-triviale machine: van data- naar procesmodellen, lost dat iets op? Chapter 2 In: Klabbers, J.H.G. [Ed], Kennen en organiseren van informatie. Over machines en actoren. [Knowledge and the organization of information. About machines and actors]. Deventer: Van Loghum Slaterus.

Campbell, D.T., and Stanley, J.C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Caplan, N. (1979). The two-communities theory and knowledge utilization. American Behavioral Scientist, 22, No. 3, 459-470.

Carley, K. (1986). An approach for relating social structure to cognitive structure. Journal of Mathematical Sociology, 12, No.2, 137-189.

Carley, M. (1980). Rational techniques in policy analysis. London: Heinemann Educational Books.

Centraal Bureau voor de statistiek, Vademecum gezondheidszorgstatistiek Nederland 1989 [ Health Care Statistics in the Netherlands]. Den Haag: Staatsuitgeverij.

Checkland; P. (1981). Systems Thinking, Systems Practice. New York: Wiley and Sons.

Churchman, C.W. (1971). The design of inquiring systems: basic concepts of systems and organization. New York: Basic Books.

Cook, T.D., and Campbell, D.T. (1979). Quasi-Experimentation. Design and Analysis Issues for Field Settings. Chicago: Rand McNally College Publishing Company.

Cowan, D. A. (1992). Adding real-time characteristics of thinking to equations of executive cognition. Proceedings of the 1992 Academy of Management Conference, 396-400.

Cuilenburg, J., van Kleinnijenhuis, J., and de Ridder, J. (1989). Tekst en betoog. Naar een computergestuurde inhoudsanalyse van betogende teksten [Text and discours. To a computer-based content-analysis of discursive texts]. Muiderberg: Coutinho.

De Geus, A.P. (1988). Planning as Learning. Harvard Business Review, 70-74.

De Man, R. (1987). Energyforecasting and the organization of the policy process: a comparative study on the use of energy forcasts in the Netherlands and the United Kingdom. Delft.

De Greene, K.B. (1982). The adaptive organization: anticipation and management of crisis. New York.

Doerner, D. (1980). On the difficulties people have in dealing with complexity. Simulation and Games, 11, No. 1, 87-106

Doerner, D., Kreuzig, H.W., Reither, F., and Stäudel, T. [Eds.]. (1983). Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität. Bern, Stuttgart, Wien: Verlag Hans Huber.

Doorn, J.A.A. van (1985). Over beleidstheorieën: Een technologische en een sociologische interpretatie [Of policy theories: a technological and a sociological interpretation]. Beleid en Maatschappij, 5, 107-116.

Douben, N.H. (1987). Economie en sociale zekerheid [Economics and Social Security], Groningen.

Dunn, W.N. (1981). Public Policy Analysis: An Introduction. New Jersey:Prentice-Hall

Dye, T.R. (1978). Understanding Public Policy. New Yersey: Englewood Cliffs.

Eden, C. (1988). Cognitive mapping. European Journal of Operational Research, 36, 1-13.

Eden, C., Smithin, T., and Wiltshire, J. (1980). Cognition Simulation and Learning. Journal of Experiential Learning and Simulation, 2, 131-143.

Felling, A.J.A. (1974). Sociaal netwerkanalyse [Social network analysis]. Alphen aan den Rijn/Brussel.

Forrester, J.W. (1961). Industrial Dynamics. Cambridge (MA): MIT Press.

Forrester, J.W. (1969). Urban Dynamics. Cambridge (MA): MIT Press.

Georgantzas, N.C., and Madu, C.N. (1990). Cognitive Processes in Technology Management and Transfer. Technological forecasting and social change, 38, 81-95.

Geurts, J.L.A. (1981). Model en spel. [Models and Games]. Phd diss, K.U. Nijmegen.

Geurts, J.L.A., G.J.P. van Griensven, J.W. Gubbels, J.A.M. Vennix (1985). The social security system in the Netherlands: an interactive simulation. Simulation and Games, 11, No. 3, 189-210.

Geurts, J.L.A., and Vennix, J.A.M. (Eds.). (1989). Verkenningen in beleidsanalyse, theorie en praktijk van modelbouw en simulatie. [Explorations in Policy Analysis, Theory amd Application of Model Building and Simulation]. Zeist: Kerckebosch.

Goote, H. (1990). Opheffing contracteerplicht en invoering maximumtarieven [Stopping the mandatory contract with suppliers of care and introduction of maximum prices]. Nederlands Tandartsenblad, 22, .

Gould, J.M. (1989a). Research Summary, Internal report.

Gould, J.M. (1989b). Transfer of System Dynamics Learning. In: P.M. Milling, and E.O.K. Zahn, [Eds]., Computer-Based Management of Complex Systems, Proceedings of the 1989 International Conference of the System Dynamics Society, Berlin: Springer,

Graham, A.K., Senge, P.M., Sterman, J.D., Morecroft, J.D.W. (1989). Computer-based case studies in management education and research. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex Systems, Proceedings of the 1989 International Conference on the System Dynamics Society, Berlin: Springer, 317-326

Greenberger, M., Crenson, M.A., and Crissey, B.L. (1976). Models in the policy process: public decision making in the computer era. New York.

Greeno, J.G. (1989). A perspective on thinking. American Psychologist, Februari, 134-141.

Grünwald, C.A. (1987). Beheersing van de gezondheidszorg [Containing the Health Care System].'s-Gravenhage.

Gubbels, J.W., Verburgh, L.D., and Heine, E.J.H., ter (1992). Van model naar beleid. Uitvoeren van beleidsexperimenten. Chapter 9 in: Post, D., and Vennix, J.A.M. [Eds]., De gezondheidszorg in model. Ons complex zorgsysteem gemeten en gewogen [The Health Care System as a Model. Our Complex Health Care System under consideration]. Houten/Zaventem: Bohn Stafleu Van Loghum.

Gusfield, J.R. (1981). Drinking-driving and the symbolic order. Chicago:University Press.

Hall, R.I. (1984). The natural logic of management policy making: its implications for the survival of an organization. Management Science, 30, No. 8, 905-927.

Hall, R.I. (1989). A training game and behavioral decision-making research tool: an alternative use of system dynamics simulation. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex Systems, Proceedings of the 1989 International Conference on the System Dynamics Society, Berlin: Springer, 221-229

Hall, R.I., Aitchison, P., Kocay, W.L., and Li, H. (1989). The organization's policy maps: types of maps, methods for soliciting and recording them, and techniques for analysis. Working document, Dep. of Business Administration, Univ. of Manitoba, Winnipeg, Canada.

Hart, S.L. (1985). Toward Quality Criteria for Collective Judgements. Organizational behavior and human decision processes, 36, 209-228.

Hart, S.L., Boroush, M., Enk, G., and Hornick, W. (1985). Through Consensus Mapping: Technology for the Structuring of Groups Decisions. Academy of Management Review, 587-600.

Helmer, O. (1966). Social Technology. New York: Basic Books.

Honigh, M. (1983). Effecten van kostenbeheersing in de gezondheidszorg. Chapter 5 in: Hoogerwerf, A., Succes en falen van overheidsbeleid [Success and Failure of Governmental Policy] Alphen aan den Rijn.

Hoogerwerf, A. (1984). Beleid berust op veronderstellingen: de beleidstheorie [Policy making is based on assumptions]. Acta Politica, 4, 493-531.

Hoogerwerf, A. (1989). De beleidstheorie uit de beleidspraktijk: Een tussenbalans. [The policy theory in practice: an in-between assessment]. Beleidswetenschap, 4, 320-341.

House, P.W. (1982). The art of public policy analysis: the arena of regulations and resources. Beverly Hills / London / New Delhi.

Huber, G.P. (1983). Cognitive style as a basis for MIS and DSS designs: much ado about nothing. Management Science, 29, No. 5, 567-582.

Isaacs, W.N., and Senge, P.M. (1992). Overcoming limits to learning in computer-based learning environments. European Journal of Operational Research, 59, 183-196.

IThink User's Guide (1991). Hannover NH

Kalff, D.J.A. (1989). Strategic decision making and simulation in Shell. In: J.J.G. Klabbers et al. [Eds]. Simulation-Gaming: On the Improvement of Competence in Dealing with Complexity, Uncertainty and Value Conflicts. Proceedings of the International Simulation and Gaming Association's 19th International Conference. , Pergamon Press.

Kaplan, R.E., Lombardo, M.M., and Mazique, M.S. (1985). A Mirror for Managers: Using Simulation To Develop Management Teams. The Journal of Behavioral Science, 21, No. 3, 241-253.

Kauffman, D.L. (1980). System One: An Introduction to Systems Thinking. Minneapolis, MN: S.A. Carlton.

Keuning, D. (1973). Algemene systeemtheorie, systeembenadering en organisatietheorie [General Systems Theory, Systems Approach and Organizational Theory]. Leiden: Stenfert Kroese.

Kim, D.H. (1989a). Learning Laboratories: designing a reflective learning environment. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex

Systems. Proceedings of the 1989 International Conference on the System Dynamics Society. Berlin: Springer, 317-326

Kim, D.H. (1989b). Computer-based case studies in management education and research. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex Systems. Proceedings of the 1989 International Conference on the System Dynamics Society. Berlin: Springer, 327-334

Kim, D.H. (1990). Total Quality and System Dynamics: Complementary Approaches to Organizational Learning. In: Proceedings of the 1990 International System Dynamics Conference. Boston, MA, 539-553.

Klabbers, J.H.G. (1972). Simulatie van een mens-machine systeem. Een systeem-psychologische benadering [Simualtion of a man-machine system, A system-psychological approach]. Phd diss. K.U.N., Nijmegen.

Klabbers, J.H.G. (1975). General system theory and social systems: a methodology for the social sciences. Nederlands Tijdschrift voor de Psychologie, 30, 493-514.

Klabbers, J.H.G. (1976). Interactieve simulatie, on-line interactie tussen mens en machine voor de bestudering van sociale systemen. In: H.A. Becker (red.), Simulatie in de sociale wetenschappen [Simulation in the Social Sciences]. Alphen aan den Rijn, 187-212.

Klabbers, J.H.G. (1977). Het modelbouwproces en het analyseren van sociale systemen. In: G. Broekstra, Knipscheer, J.S. (red.), Systemen en toekomstverkenning [Systems and Forecasting]. Leiden, 151-176.

Klabbers, J.H.G. (1983). The Use of Gaming in Futures Research and Public Policy Making. In: I. Stahl [Ed.], Operational Gaming. An International Approach. Pergamon Press.

Klabbers, J.H.G. (1985). Instruments for Planning and Policy Formation: Some Methodological Considerations. Simulations and Games, 16, 135-160.

Klabbers, J.H.G. (1986). Improvement of (self-)steering through support systems. Chapter 5 in: Sociocybernetic Paradoxes. Observation, Control and Evolution of Self-steering Systems. F. Geyer and J. van der Zouwen [Eds.]., Beverly Hills (CA): Sage

Klabbers, J.H.G. [Ed.]. (1988). Kennen en organiseren van informatie. Over machines en actoren [Knowledge and the organization of information. About machines and actors]. Deventer: Van Loghum Slaterus.

Klabbers, J.H.G. (1990). Problem Setting Through Gaming. Invited lecture at the 7th International Conference "Case Method Research and Case Method Application".

Klabbers, J.H.G. (1991). Simulation: learning environments for (self-)steering in social systems. Chapter 18 in: Autopoiesis and Configuration Theory: New Approaches to Societal Steering. R. J. In 't Veld, L. Schaap, C.J.A.M. Termeer and M.J.W. van Twist [Eds.]., Dordrecht, Boston, London: Kluwer.

Klabbers, J.H.G., and Scheper, W. (1988). Epiloog. In: Klabbers, J.H.G. [Ed], Kennen en organiseren van informatie. Over machines en actoren [Knowledge and the organization of information. About machines and actors]. Deventer: Van Loghum Slaterus.

Klein, J.H., and Cooper, D.F. (1982). Cognitive Maps of Decision-Makers in a Complex Game. Journal of Operational Research Society, 33, 63-71.

Knottnerus, J.A., and Volovics, A. (1989). Overeenstemming tussen beoordelaars [Reliability among raters]. Huisarts en wetenschap, 32, No. 2, 56-73.

Kolb, D.A. (1976). The Learning Style Inventory: Technical Manual, Boston.

Kolb, D.A. (1981). Learning style and disciplinary differences. In: A. Chickering and ass., The Modern American College, San Francisco., 232-255.

Kraan-Jetten, A. (1986). Reconstructie en evaluatie van beleidstheorieën [Reconstruction and evaluation of policy theories]. Beleid en Maatschappij, 4, 186-196.

Kramer, N.J.T.A. (1978). Systeem in probleem. Een onderzoek naar de bijdragen van de systeemleer tot de aanpak van praktijkproblemen in organisaties [Problems with systems. Research into the contribution of systems thinking to tackling organizational problems]. Leiden/Antwerpen: Stenfert Kroese.

Krippendorff, K. (1980). Content Analysis. An Introduction to Its Methodology. London: Sage.

Kroonen, A., and Bex, J.H.M. (1989). Overheidsbeleid en kostenbeheersing in de gezondheidszorg. Chapter 16 in: Reader: Organisatie gezondheidszorg en juridische aspecten [Reader: Organization Health Care System and Legal Aspects]. Nijmegen.

Kuypers, G. (1986). Beginselen van beleidsontwikkeling A [Principles of Policy Analysis A]. Muiderberg: Coutinho

Kuypers, G. (1980). Beginselen van beleidsontwikkeling B [Principles of Policy Analysis B]. Muiderberg: Coutinho

Mason R.O., and Mitroff, I.I. (1981). Challenging Strategic Planning Assumptions. Theory, Cases, and Techniques. New York, John Wiley and Sons.

Lane, D.C. (1992a). Observing the process: reflections on issue selection and model conceptualisation in a case study concerning the dynamics of programmer recruitment. System Dynamics Conference, Utrecht. 337-346.

Lane, D.C. (1992b). Modelling as Learning: A consultancy methodology for enhancing learning in management teams. European Journal of Operational Research, 59, 64-84.

Larsen, E.R., Morecroft, J.D.W., and Murphy, J. (1991). Helping Management Teams to Model: A project in the Consumer Electronics Industry. System Dynamics Conference, Utrecht, 281-290.

Leeuw, F.L. de, (1983). Bevolkingsbeleid en reproduktief gedrag: een studie naar de theorieën over gedrag die aan beleid ten grondslag liggen [Population policy making and reproductive behaviour: a study of the theories regarding the behaviour underlying policy making]. Phd diss., R.U. Leiden.

Leeuw, F.L. de, (1986). Beleidstheoretisch onderzoek toen en thans [Policy-theoretical research then and now]. Beleid en Maatschappij, 1, 27-39.

Maesen, L.J.G. van der (1987). Transformatie van de gezondheidszorg in Nederland tussen 1974 en 1987 [Transformation of the Dutch Health Care System in between 1974 and 1987]. Assen/Maastricht: Van Gorcum.

Mason, R.D., and Mitroff, I.I. (1981). Challenging Strategic Planning Assumptions. Theory, Cases, and Techniques. New York: Wiley and Sons.

Meadows, D.H. (1980). The Unavoidable A Priori. Chapter 2 in: Elements of the System Dynamics Method, J Randers (Ed.), Cambridge, MA: The MIT Press.

Meadows, D.H., Richardson, J., and Bruckmann, G. (1982). Groping in the dark: the first decade of global modelling. Chicester/New York.

Meadows, D.H., and Robinson, J.M. (1985). The electronic oracle: computer models and social decisions. Chistester/New York.

Mitchell, D.E. (1980). Social science impact on legislative decisionmaking: process and substance. Educational Researcher 9, 10.

Mokking, H.G.A. (1986). Ziekenfondscijfers als parameters voor het handelen van huisartsen [Health Care Insurance company statistics as parameters for the medical behaviour of general practitioners]. Meppel: Krips.

Montazemi, A.R., and Conrath, D.W. (1986). The Use of Cognitive Mapping for Information Requirements Analysis. MIS Quarterly, March 1986, 45-56.

Montfort, A.P.W.P., and Vandermeulen, L. (1988). Is de gezondheidszorg echt duurder geworden? [Has the health care system indeed become more expensive?]. Medisch Contact, 43, 1155-1158

Morecroft, J.D.W. (1988). System Dynamics and microworlds for policymakers. In: European Journal of Operational Research, 35, 301-320.

Morecroft, J.D.W. (1992). Executive knowledge, models and learning. European Journal of Operational Research, 59, 9-27.

Morecroft, J.D.W., Lane, D.C., and Viita, P.S. (1989). Modelling a Biotechnology Startup Firm. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex Systems. Proceedings of the 1989 International Conference on the System Dynamics Society. Berlin: Springer, 157-176

Morgan, G. (1986). Images of organization. Beverly Hills, CA: Sage.

Muskens, G.J. (1980). Frames of meaning, are they measurable? A methodological critique of the content analysis of illustrated periodical magazines. Phd diss. K.U.N., Nijmegen:

Nationaal Ziekenhuisinstituut, De intramurale gezondheidszorg in cijfers per 1 januari 1986 [Health Care Statistics regarding in-patient care. Figures from january 1st 1986].

Nationaal Ziekenhuisinstituut, De intramurale gezondheidszorg in cijfers per 1 januari 1988 [Health Care Statistics regarding in-patient care. Figures from january 1st 1988].

Nelissen, N., Geurts, J., and Wit, de H. (1986). Het verkennen van beleidsproblemen (Exploring Policy Problems). Zeist: Kerckebosch.

Osgood, C.E. Sporta, S., and Nunnally, J.C. (1956). Evaluative assertion analysis., Litera, 3, 47-102.

Oud, J.H.L. (1978). System-methodologie in sociaal-wetenschappelijk onderzoek [Systems Methodology in Social Sciences Research]. Phd diss. K.U.N., Nijmegen: Alpha.

Plaats, J.J. van der, and Bex, J.H.M. (1989). Een historisch overzicht van de ontwikkeling van de gezondheidszorg. Chapter 1 in: Reader: organisatie gezondheidszorg en juridische aspecten [Reader: Organization Health Care System and Legal Aspects]. Nijmegen.

Popping, R. (1983). Overeenstemming en associatie in sociologisch onderzoek [Reliability and association in sociological reseach]. Mens en Maatschappij, Jrg. 58, No. 2, 132-150.

Post, d. (1984a). Wijzigingen in prescriptie. Het effect van de oktober- en februarimaatregel [Changes in Prescription. The effects of the october and february measures]. Medisch Contact, 39, 19-21

Post, D. (1984b). Verwijzen en prescriptie. Een paar apart [Referrals and Prescriptions. Of a different kind]. Zwolle: Regionaal Ziekenfonds Zwolle.

Post, D. (1989). Gezondheidszorg op rantsoen [Health Care System on budget]. Kampen: Kok.

Post, D., [Ed] (1990). Hier meer daar minder. Een onderzoek naar het gebruik van gezondheidsvoorzieningen in twee ziekenfondsgebieden met het oog op instrumentontwikkeling in het kader van externe kwaliteitstoetsing [More here and less there. A study into the use of health care provisions in two health care insurance regions in light of the development of a tool to assess the external quality]. Meppel: Ziekenfonds Het Groene Land.

Post, D., and Vennix, J.A.M. [Eds]., (1992). De gezondheidszorg in model. Ons complex zorgsysteem gemeten en gewogen [The Health Care System as a Model. Our Complex Health Care System under consideration]. Houten/Zaventem: Bohn Stafleu Van Loghum.

Prawat, R.S. (1989). Promoting Access to Knowledge, Strategy, and Disposition in Students: A Research Synthesis. Review of Educational Research, 59, No. 1, 1-41.

Premkumar (1989). A cognitive study of the decision-making process in a business context: Implications for design of expert systems. International Journal of Man-Machine Studies, 31, No. 5, 557-572.

Quade, E.S. (1989). Analysis for Public Decisions. New York: North-Holland.

Raiffa, H. (1968). Decision Analysis. Reading, Mass.: Addison-Wesley.

Randers, J. (1974). System Dynamics Modelling. Textbook for the 1974 system dynamics modelling course.

Randers, J. (Ed.) (1980). Elements of the System Dynamics method. Cambridge, MA.

Ree, C.M. van der, Mokking, H.G.A., Post, D., and Gubbels, J. (1991). De werkbelasting van de huisarts [The workload of general practitioners]. Medisch Contact, 46, 465

Reiman, B.C., and Ramanujam, V. (1992). Acting Versus Thinking: A Debate Between Tom Peters and Michael Porter. Planning Review, March/April, 36-43

Rein, M. (1976). Social Science and Public Policy. Pinguin Education.

Reinhardt, R., and Schwelker, U. (1992). Building learning organizations: implementing organizational learning processes, and the theory of self-referential systems. System dynamics conference, Utrecht, 573-584.

Richardson, G.P., Andersen, D.F., Rohrbaugh, J., and Steinhurst, W. (1992). Group Model Building. Proceedings System Dynamics, 595-604.

Richardson, G.P., Vennix, J.A.M., Andersen, D.F., Rohrbaugh, J., and Wallace, W.A. (1989). Eliciting Group Knowledge for Model-Building. In: P.M. Milling, E.O.K. Zahn (Eds.), Computer Based Management of Complex Systems, Proceedings of the 1989 International Conference on the System Dynamics Society, Berlin: Springer, 343-357

Richardson, G.P., and Pugh, A.L. (1981). Introduction to system dynamics modelling with DYNAMO, Cambridge, MA.

Ridder, J. de (1990). Informatie zoeken in politieke teksten [Searching for information in political texts]. Massacommunicatie, 18, 117-136.

Ringeling, A.B. (1985). Beleidstheorieën en theorieën over beleid [Policy-theories and theories about policy]. Beleid en Maatschappij, 11, 275-284.

Roberts, N., Andersen, D.F., Deal, D., Garet, M., and Schaffer, W. (1983). Introduction to computer simulation: A system dynamics modeling approach, Reading, MA.

Roscam Abbing, E.W. (1989). De plannen van de commissie Dekker en hun uitvoering. Chapter 19 in: Reader: Organisatie gezondheidszorg en juridische aspecten [Reader: Organization Health Care System and Legal Aspects]. Nijmegen.

Rosenthal, U., Schendelen, M.P.C.M., van and Ringeling, A.B. (1986). Openbaar bestuur: organisatie, beleid en politieke omgeving [Public policy: organization, policy and political context]. Alphen aan den Rijn: Tjeenk Willink.

Rutges, A. (1976). Begrippen rond simulatie. In: Becker, H.A. [Eds.] Simulatie in de sociale wetenschappen [Simulation in the Social Sciences]. Alphen aan den Rijn: Samsom.

Schaik, F.D.J., van (1988). Effectiveness of decision support systems, Phd diss. T.U. Delft.

Schein, E.H. (1993). How Can Organizations Learn Faster? The Challenge of Entering the Green Room. Sloan Management Review/Winter, 85-92.

Schellekens, W. (1989). Dekker en kwaliteit gedekt? [Dekker and quality assured?]. Medisch Contact, 13, 427.

Scheper, W.J. (1991). Group Decision Support Systems. An inquiry into theoretical and philosophical issues. Phd diss. R.U.Utrecht.

Schieber, G.J., and Poullier, J.P. (1988). International Health Spending and Utilization Trends. Health Affair, 7, 105-112.

Scholten, J.H. (1990). Vragen aan Simons [Questions to Simons]. Nederlands Tandartsenblad, 45.

Schön, D. (1983). The Reflective Practitioner. New York: Basic Books.

Schrijver, G. (1987). Regionalisatie en competitie in de gezondheidszorg. Het rapport van de commissie-Dekker [Regionalization and competition in the Health Care System. The report of the Dekker-committee]. ESB 15/22-4, 368-371.

Segers, J.G.H., and Hagenaars, J.A.P. [Eds] (1980). Sociologische onderzoeksmethoden, Deel II. Technieken van causale analyse. [Sociological research methods. Part II. Techniques for causal analyses]. Assen: Van Gorcum.

Senge, P.M. (1992). You Can't Get There from Here: Why Systems Thinking is Inseparable from Learning Organizations. Proceedings System Dynamics, Utrecht, 675-684.

SIG, LMR jaarboek 1985, 1987 [ LMR annual book 1985, 1987]. Utrecht: SIG.

SIG, Regio-atlas 1986 [Regional map, 1986]. Utrecht: SIG.

Sleegers, P.J.C. (1991). School en beleidsvoering. Een onderzoek naar de relatie tussen het beleidsvoerend vermogen van scholen en het benutten van de beleidsruimte door scholen [School and policy making. A study into the relationship between the policy making abilities of schools and the use of policy opportunities by schools]. Phd diss. K.U.N, Nijmegen.

Speld, G.D.J. van der, (1990). Regionale verschillen in gebruik ziekenhuisvoorzieningen [Regional differences in the use of hospital provisions]. HZH, 4, 136-139.

Stenhouse, L. (1975). Introduction to Curriculum Research and Development. London: Heinemann.

Stichting KLOZ informatiesysteem gezondheidszorg, KISG86, Jaarboek 1986 [KISG86 Annual Book 1986].,Noordwijk: Stichting KISG.

Stone, D.A. (1987). Tales of America. New York: Times Books

Strauch, R. (1975). 'Squishy' problems and Quantitative Methods. Policy Sciences, June, 175-184.

Structuur en financiering gezondheidszorg: nota Verandering verzekerd: stapsgewijs op weg naar een nieuw stelsel van zorg [Structure of and financing the Health Care System: Report 'Change Insured: a step-wise introduction of a new system']. Den Haag: Staatsuitgeverij (1988).

Tennyson, R.D., Thurlow, R., and Breuer, K. (1988). Problem-Oriented Simulations to Develop and Improve Higher-Order Thinking Strategies. Computers in Human Behavior, 3, 151-165.

Theil, H. (1970). On the Estimation of Relationhips Involving Qualitative Variables. American Journal of Sociology, 76, 103-154.

Thissen, W. (1984). Systeembenadering: een leidraad bij beleidsanalyses [Systems approach: a guide for policy analyses]. Systemica, 3, No. 4. 342-347.

Toen en nu. Ziekenfondsen en VNZ [The past and the present. Regionale Health Care Insurance Companies and VNZ]., 1984.

Toval, A., and Flores, M. (1987). Computer Systems Simulation in Education: Description of an Experience. Computer Education, 11, No. 4, 293-303.

Vademicum Gezondheidszorgstatistiek Nederland [ Vademicum Health Care Statistics The Netherlands]., 1989.

VanGundy, A.B. (1988). Techniques of structured problem solving. New York: Van Nostrand Reinhold Company.

Ven, W.P.M.M. van de (1989). Overheid of markt [Government or market]. Inzet, 12, 12-14

Vennix, J.A.M. (1990). Mental models and computer models. Design and evaluation of a computer-based learning environment for policy-making. Phd diss. K.U.N., Nijmegen.

Vennix, J.A.M. (1992). Beslissingsondersteuning voor slecht gestructureerde problemen: modelbouw en computersimulatie met behulp van STELLA. Goed geinformeerd. Recente ontwikkelingen en toepassingen van informatietechnologie in de sociale wetenschappen. [Well-informed. Recent developments and applications of information-technology in the Social Sciences]. Boekaflevering Mens and Maatschappij, 66e jaargang., F. Popping and J.L. Peschar [Eds.]. 83-98.

Vennix, J.A.M., and Geurts, J.L.A. (1987). Communicating insights from complex simulation models: a gaming approach. Simulation and Games, 18, No. 3, 321-343.

Vennix, J.A.M., and Gubbels, J.W. (1988). 'Workshop'-benadering. De kostenbeheersing in de gezondheidszorg. Chapter 17 in: W. Derksen, A.F.A. Korsen, and A.F.M. Bertrand [Eds.], De praktijk van onderzoek. Problemen bij onderzoek van politiek, bestuur en beleid [Research in practice. Problems encountered at the research of politics, public policy and policy making].

Vennix, J.A.M., and Gubbels, J. (1992). Knowledge elicitation in conceptual model building: A case study in modeling a regional Dutch health care system. European Journal of Operational Research, 59, 85-101.

Vennix, J.A.M., Gubbels, J.W., and Post, D. (1986a). Spreekuurbezoek, verwijzen en opnemen als proces, I: Patienten onder behandeling van de huisarts [Consultations, referrals and admissions into hospitals as a proces, I: Patients under treatment by a general practitioner]. Medisch Contact, 27, 852-854.

Vennix, J.A.M., Gubbels, J.W., and Post, D. (1986b). Spreekuurbezoek, verwijzen en opnemen als proces, II: Patienten onder behandeling van de specialist en ziekenhuisopnames [Consultations, referrals and admissions into hospitals as a proces, II: Patients under treatment by a medical specialist and admissions into hospital]. Medisch Contact, 28, 895-898.

Vennix, J.A.M., Gubbels, J.W., Post, D., and Poppen, H. (1988). A Structured Approach to Knowledge Acquisition in Model Development. Proceedings of the International System Dynamics Conference, San Diego (CA).

Vennix, J.A.M., and Scheper, W.J. (1990). Modelling as Organizational Learning: an Empirical Perspective. Proceedings of the 1990 System Dynamics Conference, Boston (MA), 1199-1210.

Vennix, J.A.M., Schmeets, J.J.G., and Wester, F. (1987). Beleidstheorieën en simulatiemodellen: De relatie tussen het sociale zekerheidsstelsel en de Nederlandse economie [Policy-theories and simulation models: The relationship between the social security system and the Dutch economy]. Beleidswetenschap, 3, 249-271.

Vennix, J.A.M., Smits, R.E.H.M., and Geurts, J.L.A. (1989). De functie van computersimulatie: inleiding in de systeemdynamica. Chapter 3 in: Geurts, J., and Vennix, J. [Eds]., Verkenningen in beleidsanalyse, theorie en praktijk van modelbouw en simulatie [Explorations in Policy Analysis, Theory and Application of Model Building and Simulation].Zeist: Kerckebosch.

Vennix, J.A.M., Verburgh, L.D., Gubbels, J.W., and Post, D. (1990). Eliciting group knowledge in a computer-based learning environment. Proceedings of the 1990 System Dynamics Conference, Boston (MA), 1187-1198.

Verburgh, L.D., Gubbels, J.W., Vennix, J.A.M., and Post, D. (1990). Model-based Analyses of the Dutch Health Care System. Proceedings of the 1990 System Dynamics Conference, Boston (MA), 1211-1225.

Verburgh, L.D., and Gubbels, J.W. (1992). Van conceptueel model tot computermodel. Chapter 8 in: Post, D., and Vennix, J.A.M. [Eds]., De gezondheidszorg in model. Ons complex zorgsysteem gemeten en gewogen [The Health Care System as a Model. Our Complex Health Care System under consideration]. Houten/Zaventem: Bohn Stafleu Van Loghum.

Vickers, G. (1965). The Art of Judgement: AStudy of Policy Making. London: Chapman and Hall.

Vickers, G. (1972). Commonly ignored elements in policy-making. Policy Sciences, 3.

Von Foerster, H. (1984). Principles of Self-Organization - In a Socio-Managerial Context. In: Self-Organization and Management of Social Systems, Insights, Promises, Doubts, and Questions, H. Ulrich and G.J.B. Probst [Eds]., Berlin Heidelberg New York Tokyo: Springer-Verlag.

Waern, Y. (1990). On the dynamics of mental models. In: Mental models and Human-Computer Interaction 1. (D.Ackermann and M.J. Tauber (Eds.). Elsevier (North-Holland).

Watt, C.H. (1977). Why won't anyone believe us? Simulation, 1-3

Webber, M.M., and Rittel, H.W.J. (1973). Dilemmas in a General Theory of Planning. Policy Sciences, 4, No. 2, 155-169.

Weick, K.E. (1979). Cognitive Processes in Organizations. Research in Organization Behavior, 1, B.M. Staw [Ed], Greenwich, Connecticut: JAI Press

Weick, K.E., and Bougon, M.G. (1986). Organizations as cognitive maps: charting ways to success and failure. In: Sims, H.P., D.A. Gioia [Eds.], The thinking organization. San Francisco/London.

Weil, H.B. (1980). The evolution of an approach for achieving implemented results from system dynamics projects. In: Randers, J. (Ed)., Elements of the system dynamics method. Cambridge, MA.

Wijnberg, B. (1989). Terugtredende overheid, meer aan het veld overlaten [Withdrawing Central Government, more is left to the field itself]. Nederlands tandartsenblad, 16.

Zitman, G.T. (1989). Een poging om 'Dekker' te redden [An attempt to save 'Dekker']. Medisch Contact, 38, 1207-1210.

Zouwen, J. van der, (1988). Onderzoek naar methoden van waarnemen. Hoe de zelf-referentie van een sociaal systeem de (methoden)onderzoeker parten speelt. Chapter 4 In: Klabbers, J.H.G. [Ed], Kennen en organiseren van informatie. Over machines en actoren [Knowledge and the organization of information. About machines and actors]. Deventer: Van Loghum Slaterus.

# SAMENVATTING

## PARTICIPATIEVE MODELBOUW EN BELEIDSONDERSTEUNING

Het ontwikkelen en uitvoeren van beleid binnen de complexiteit en dynamiek waar organisaties vandaag de dag mee geconfronteerd worden, is een niet gemakkelijke opgave. Problemen en ontwikkelingen lijken in grote mate met elkaar samen te hangen, veranderingen volgen elkaar steeds sneller op en het aantal betrokken actoren (partijen) is vaak niet gering. De vraag is op welke wijze binnen dit geheel aan inhoudelijke en sociale complexiteit en dynamiek, adequaat handelen tot stand kan worden gebracht.

Een methode die bij uitstek geschikt lijkt om binnen bovenstaande complexiteit en dynamiek tot verstandig en weloverwogen beleid te komen, is de participatieve modelbouw. De participatieve modelbouw biedt de betrokken actoren de gelegenheid om op voorgestructureerde wijze een modelbouwtraject te doorlopen met behulp waarvan inzicht in de complexiteit kan worden verkregen, nieuwe informatie kan worden verwerkt, toekomstverkenningen kunnen worden uitgevoerd, en consensus omtrent het te voeren beleid tot stand kan worden gebracht. Het gaat hierbij om de constructie van een beeld (model) van de situatie of het probleem waar de betrokkenen zelf zoveel mogelijk in participeren: de betrokkenen construeren hun eigen (gedeelde) afbeelding van de situatie met behulp waarvan (gedeeld) inzicht in deze situatie verkregen wordt en tot handelen kan worden overgegaan.

Het *samen* met opdrachtgevers en betrokkenen ontwikkelen en gebruiken van (computer) modellen [participatieve modelbouw] lijkt een veelbelovende vorm van beleidsondersteuning te zijn, omdat het de kwaliteit van het model vergroot en in de hand werkt dat de uitkomsten van de analyses geaccepteerd en geimplementeerd worden. Daar het bij de constructie van het model om de constructie van een gemeenschappelijk beeld gaat, is participatieve modelbouw met name geschikt voor het leveren van ondersteuning bij problemen waarvoor een helder en gedeeld model of beeld van de situatie ontbreekt: de zogenaamde *slechtgestructureerde* (beleids)problemen. Dit zijn problemen die gekenmerkt worden door een groot aantal betrokkenen, conflicterende waarden, onzekerheid en een schier oneindig aantal mogelijke oplossingen.

Bij het ontwikkelen en analyseren van het door de betrokkenen geconstrueerde en gevalideerde model van het beleidsprobleem in kwestie, biedt de participatieve modelbouw ondersteuning aan de zogenaamde *beleidsontwikkelingsfasen*. Dat wil zeggen, ze assisteert in het definiëren en structureren van het probleem, het in kaart brengen van mogelijke alternatieve oplossingen en hun oplossingskracht (effecten), om tot concrete aanbevelingen te (kunnen) komen. Implementatie en evaluatie van de ingevoerde maatregelen vallen daarmee buiten de primaire scope van de door participatieve modelbouw te bieden ondersteuning.

## EVALUATIE VAN DE DOOR PARTICIPATIEVE MODELBOUW GELEVERDE BIJDRAGE

Het doel van deze studie is om in kaart te brengen of en in welke mate participatieve modelbouw in staat is om bovenstaande claims waar te maken. Zij dient inzichtelijk te maken in hoeverre de methode een bijdrage kan leveren aan het omgaan met en het oplossen van slechtgestructureerde beleidsproblemen.

De evaluatie van de effecten van participative modelbouw is tot nu toe overwegend anecdotisch van aard geweest. Met verve is beschreven hoe lovend deelnemers zich over de methode hebben uitgelaten (interessant, waardevol, voor herhaling vatbaar etc.). Echter systematisch empirisch onderzoek heeft amper plaatsgevonden. Een uitzondering hierop is het door Vennix (1990) uitgevoerde onderzoek beschreven in zijn proefschrift 'Mental models and computer models. Design and evaluation of a computer-based learning environment for policy-making'. In de opzet en uitvoering van het huidige onderzoek heeft het door Vennix uitgevoerde onderzoek dan ook als referentiekader gediend. In het onderstaande zal een beschrijving worden gegeven van de manier waarop in de huidige studie getracht is de effecten van participatieve modelbouw in kaart te brengen. In deze beschrijving zullen achtereenvolgend aan bod komen: de onderzoeks-doelstellingen en -vragen die door het onderzoek beantwoord dienen te worden, de gehanteerde criteria ter beoordeling van de door de methode te leveren bijdrage, de gebruikte meetinstrumenten, het aangeboden programma, de onderzoeksgroep (populatie) en de gehanteerde onderzoeksopzet (het design).

*» onderzoeksdoelstellingen and -vragen: wat te onderzoeken*

Hoewel zowel de huidige studie als het door Vennix uitgevoerde onderzoek zich ten doel stellen de conceptuele bijdrage of impact van participatieve modelbouw empirisch vast te stellen (gaat men door deelname aan het programma anders denken, het probleem anders definiëren, krijgt men meer kennis van de ingewikkelde materie), richt de huidige studie zich niet alleen op de veranderingen die plaatsvinden bij de *individuele* deelnemers (verandert hun denken over het probleem - verandert hun mentale map), maar ook op de veranderingen die plaatsvinden bij de deelnemers als groep: de inter-individuele veranderingen. Dit omdat de richting van het collectieve handelen (de organisatie, de groep) in grote mate bepaald wordt door de wijze waarop de betrokkenen er te zamen over denken. De te beantwoorden hoofdvragen zijn dan ook:

- verandert participative modelbouw de mentale mappen van *individuele* deelnemers?
- resulteert participatieve modelbouw in een toename in de *homogeniteit* van de mentale mappen van hen die er aan deelnemen?

De eerste vraag betreft een individuele verschuiving in het denken over het probleem. De richting waarin verschoven wordt is hierbij van ondergeschikt belang aan de verandering zelf. Onderzocht wordt of de deelnemers andere concepten zijn gaan gebruiken en andere verbanden zijn gaan leggen door deelname aan het programma. De tweede vraag daarentegen, heeft betrekking op de richting waarin de verandering, in de eerste vraag beschreven, gaat plaatsvinden. Het is de verwachting dat deelname tot gevolg heeft dat men meer op een overeenkomstige wijze gaat nadenken over het probleem. De overeenkomst in de wijze waarop heeft dan betrekking op zowel de structuur (strategische kennis) als inhoud (domein-specifieke kennis) van het denken.

Door het onderscheid domein-specifiek en strategisch toe te passen op het onderscheid tussen individuele en inter-individuele verandering kunnen de volgende vier onderzoeksvragen worden onderscheiden:

| | individueel | inter-individueel |
|---|---|---|
| strategisch | leidt deelname tot hogere scores op de onderscheiden criteria? (I) | leidt deelname tot afname in de variantie bij minimaal gelijkblijvende gemiddelden voor de onderscheiden criteria ? (III) |
| domein-specifiek | leidt deelname tot hogere scores op de onderscheiden criteria? (II) | leidt deelname tot afname in de variantie bij minimaal gelijkblijvende gemiddelden voor de onderscheiden criteria ? (IV) |

In aanvulling op de bovenstaande vier vragen, stelt het onderzoek zich ook ten doel inzicht te krijgen in de mate waarin de deelnemers door deelname bewust worden van hoe anderen over het betreffende probleem denken. Dit omdat het de verwachting is dat bewustzijn van elkaars opvattingen en kennis een noodzakelijke voorwaarde is om wederzijdse beinvloeding (en dientengevolge homogenisering van het denken) tot stand te brengen. De vijfde onderzoeksvraag luidt dan ook:

(V): Leidt participatieve modelbouw tot een toename in de mate waarin men bewust is (kennis heeft) van elkaars opvattingen over het probleem?

*» criteria ter beoordeling van verandering*

Nu de belangrijkste onderzoeksvragen zijn samengevat, moet worden aangeven op basis waarvan (welke criteria) kan worden vastgesteld of verandering in domein-specifieke en strategische kennis op individueel en inter-individueel niveau heeft plaatsgevonden.

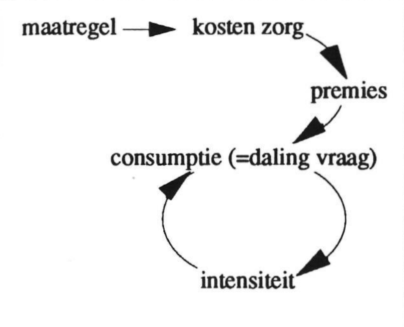
In tegenstelling tot de Vennix studie, zijn de in de huidige studie gebruikte criteria niet afkomstig uit de 'wereld van beleidstheorieën'. Gezien de beperkte hoeveelheid door Vennix gevonden verschillen tussen voor- en nameting en de discussie die nog plaatsvindt met betrekking tot wat een beleidstheorie tot een ' goede' beleidstheorie maakt, lijkt het ons wenselijker om dichter bij de inhoud en aard van de interventie te blijven en verandering te definiëren in termen van het aangeboden programma: de participatieve systeem dynamische modelbouw. De criteria ter beoordeling van verschuivingen in strategische en domein-specifieke kennis zijn dan ook rechtstreeks afkomstig uit de systeem dynamica (de indeling strategisch / domein-specifiek is afkomstig uit de educationele psychologie en door Klabbers (1990) geïntroduceerd in de 'wereld van de beleidsondersteuning'):

- *Strategische kennis* bestaat uit de systeem dynamica elementen 'structuur', 'amplificatie' en 'vertraging'. Door meer van één of meer van deze elementen in het denken over het probleem op te nemen, verhoogt men zijn of haar score op de strategische kennis dimensie. De inter-individuele score neemt toe in het geval men meer in dezelfde mate van deze elementen gebruik gaat maken. Bijvoorbeeld door allen vertragingen in het denken over het probleem op te nemen in plaats van slechts een enkeling.

- *Domein-specifieke kennis* betreft de kennis over het domein dat door het specifieke probleem wordt omspannen. Het gaat hierbij om de inhoud van het denken over het probleem. Dientengevolge is voor ieder probleem de domein-specifieke kennis anders. Om de mate van domein-specifieke kennis over het probleem in kaart te brengen kan gebruik worden gemaakt van de systeem dynamica elementen 'exogeniteit', 'endogeniteit' en 'multidisciplinariteit'. Daarnaast kan ook de 'mate van overeenkomst met het geconstrueerde externe model' gebruikt worden om uitspraken te doen over de inhoud van het denken van de deelnemers over het probleem.

*» meetinstrumenten: hoe een verandering in denken vast te stellen*

Nu we de criteria hebben vastgesteld met behulp waarvan we theoretisch in kaart kunnen brengen wat er in het denken over het probleem (de conceptualisering van het probleem) verandert, dient te worden aangegeven hoe we het denken van de deelnemers over het probleem willen vaststellen (meten). In navolging van Vennix, gebruiken we hiertoe een voor- en nameting waarin de deelnemers gevraagd worden antwoord te geven op open vragen die betrekking hebben op het probleem(veld). Door de antwoorden op een bepaalde manier te analyseren (met behulp van de zogenaamde 'cognitive mapping' methode), zijn we in staat om uit de teksten (antwoorden) mentale mappen (concepten en relaties tussen deze concepten) af te leiden die vervolgens geanalyseerd kunnen worden. De mentale mappen kunnen als een handige manier van beschrijving van de teksten als neerslag van het denken worden beschouwd. Een dergelijke mentale map en de bijbehorende tekst ziet er bijvoorbeeld als volgt uit:




Tekst (fictief):
De maatregel heeft tot gevolg dat de kosten van de zorg zullen toenemen. Hierdoor zullen de premies moeten toenemen, wat op termijn remmend zal gaan werken op de consumptie. Echter, de daling van de vraag zal gecompenseerd worden door een toename van de intensiteit van de zorg waardoor de totale consumptie amper zal dalen.

Naast de voor- en nameting, dienen de deelnemers ook na afloop van iedere sessie een aantal vragen te beantwoorden. Deze vragen hebben betrekking op zowel de mate waarin de bijeenkomst een bijdrage aan de strategische en domein-specifieke kennis heeft geleverd als ook hoe interessant, makkelijk, zinvol, lang en leerzaam ze de bijeenkomst hebben gevonden. Voor een overzicht van de voor- en nameting en de vragenlijsten wordt verwezen naar bijlagen 1 tot en met 5.

*» het aangeboden programma: participatieve modelbouw*

Participatieve modelbouw kan als een participatieve variant van systeem dynamische modelbouw worden beschouwd. Dit betekent dat de fasen die bij systeem dynamische modelbouw worden doorlopen (conceptualisering van het probleem in een kwalitatief model, formalisering van het probleem in een kwantitatief model, analyse van het kwantitatieve model en uitvoering van beleidsexperimenten) zo veel mogelijk sámen met de klant worden uitgevoerd. In deze studie is de participatieve modelbouw toegepast binnen de gezondheidszorg. In de toepassing heeft de voordurende stijging van de kosten van de zorg daarbij centraal gestaan. In het in het kader van deze studie aangeboden programma is van zogenaamde werkboeken gebruik gemaakt. Werkboeken bestaan uit teksten die door de deelnemers, ter voorbereiding op een bijeenkomst, dienen te worden doorgenomen. Zij bevatten informatie, oefeningen en vragen die betrekking hebben op meningen van de deelnemers. Door van werkboeken gebruik te maken kan in relatief korte tijd door onervaren modelbouwers onder begeleiding van ervaren modelbouwers het systeem dynamische modelbouwtraject participatief worden doorlopen. Het concrete in deze studie aangeboden programma ziet er als volgt uit:

- bijeenkomst 1: structuur van het model
  In deze bijeenkomst ligt de nadruk op de elementen die te zamen het probleem(veld) opspannen en verantwoordelijk zijn voor het probleem van de stijgende kosten in de gezondheidszorg. De bijeenkomst duurt ongeveer vier uur en bestaat uit afwisselend plenaire en subgroepsactiviteiten. De bedoeling is dat aan het einde van de bijeenkomst een door de deelnemers geaccordeerd en gedeeld model van het probleem is ontstaan. Om de discussie richting te geven en te versnellen wordt gebruik gemaakt van een door de onderzoeker gemaakt 'voorlopig' model dat tijdens de bijeenkomst door de deelnemers gewijzigd kan worden.

- bijeenkomst 2: exogene invloeden
  In de tweede bijeenkomst ligt de nadruk op het in kaart brengen en bestuderen van externe invloeden die wel van invloed zijn op het probleem maar in verhouding niet of nauwelijks door het probleem worden beïnvloed. Te denken valt dan bijvoorbeeld aan de invloed van vergrijzing of medische technologie, of de toename van het aantal aanbieders van zorg. Deelnemers moeten hun verwachtingen uitspreken ten aanzien

van de ontwikkeling van elk van deze mogelijke invloeden en de relatie tot het model beschrijven (waar en hoe(veel) grijpt het aan?), Vervolgens kunnen dan de dynamische gevolgen van de gedeelde assumpties bestudeerd worden. Dit kan leiden tot óf een aanpassing van het individuele denken óf een aanpassing van het (externe) model.

- Bijeenkomst 3: Beleidsexperimenten
  In deze bijeenkomst worden mogelijke maatregelen geselecteerd en aan het model gekoppeld waardoor de dynamische effecten van de maatregelen bestudeerd kunnen worden. Dit om tot keuze van een of meerdere maatregelen te komen voor een mogelijk in te gaan implementatietraject.

- Bijeenkomst 4: Conclusies en afsluiting
  In deze bijeenkomst wordt terug gekeken op de bevindingen uit eerdere bijeenkomsten. De resultaten worden gekoppeld aan het reeds bestaande beleidsplan: het plan wordt aangevuld en/of gewijzigd zodat ook daadwerkelijk van de verkregen inzichten gebruik gaat worden gemaakt.

De lezer die meer inzicht wil krijgen in de hoe het model van de Nederlandse gezondheidszorg er uit ziet en op welke wijze er analyses mee kunnen worden uitgevoerd, wordt verwezen naar hoofdstuk 5. Daar wordt onder andere de structuur van het basismodel uiteengezet, wordt de invloed van de groei van het aantal aanbieders van zorg en bevolking beschreven en wordt aangegeven wat het effect van een bepaalde maatregel is.

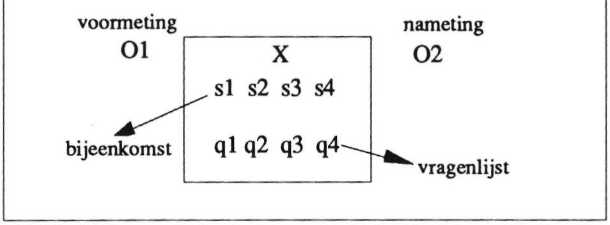
*» onderzoeksgroep (populatie)*

Teneinde de effecten van participatieve modelbouw empirisch te kunnen bestuderen, is het noodzakelijk dat zij waar mogelijk wordt toegepast in een reeële situatie om de representativiteit (externe validiteit) van de bevindingen te vergroten. In onze studie is een regionaal ziekenfonds bereid gevonden aan zowel de participatieve modelbouw als het er aan verbonden onderzoek deel te nemen. Het ziekenfonds was tot deelname bereid omdat het zich wilde voorbereiden op de verwachte decentralisatie en verzelfstandiging (ingezet door de plannen van de Commissie Dekker). Daarnaast zag zij het programma als een gelegenheid om de afdelingen Overeenkomst and Tarieven en Medische Dienst conceptueel te integreren - de denkkaders van de betrokkenen op elkaar af te stemmen.

In totaal hebben 22 mensen aan het onderzoek deelgenomen (de 'intention to treat' groep. Van deze 22 hebben er 18 in voldoende mate aan het programma deelgenomen om zich te kunnen kwalificeren voor de 'on-treatment' groep. Van de deelnemers is informatie ingewonnen (vragenlijst) met betrekking tot hun leeftijd, opleidingsniveau, specialisatie (niet medisch), het aantal jaren bij de organisatie, de mate van ervaring met beleid en beleidsontwikkeling, de afdeling waar men werkzaam is, het geslacht en de organisatie (fusiepartner) waar men oorspronkelijk uit afkomstig was (de organisatie was kortgeleden gefuseerd). Deze informatie is gebruikt bij het specificeren van de relatie tussen participatieve modelbouw en opgetreden effecten. Bijvoorbeeld door te onderzoeken of er bij deelnemers uit de Medische Dienst andere effecten optreden dan bij deelnemers uit de Overeenkomst and Tarieven afdeling. Voor een overzicht van de scores op de bovengenoemde vragen wordt verwezen naar hoofdstuk 5, paragraaf 5.3. Tevens is in hoofdstuk 5 een beschrijving gegeven van de positie van ziekenfondsen in het gezondheidszorgstelsel voor zoverre relevant voor de ontwikkeling en beschrijving van het model van de gezondheidszorg (paragraaf 5.2).

*» onderzoeksopzet (design)*

Nu is uiteengezet welke organisatie en individuele deelnemers aan de participatieve modelbouw hebben deelgenomen teneinde de onderzoeksvragen te beantwoorden, dient de opzet van het onderzoek, waarbinnen aan het bovenstaande uitvoering kan worden gegeven, te worden uiteengezet. Om het verschil in denken (neergeslagen in een voor- en nameting) te kunnen bestuderen is gekozen voor het zogenaamde 'One Group Pretest Posttest Design'. Van belang is hierbij te vermelden dat het design dat wij voor ogen hadden ('Untreated Control Group Design With Pretest and Posttest') om verschillende redenen niet kon worden uitgevoerd (zie ook hoofdstuk 4). Het gebruikte design kan als volgt worden weergegeven:



One-Group Pretest Posttest Design (Campbell & Stanley, 1966)

Uit de figuur blijkt dat er in deze studie voor gekozen is om de deelnemers voorafgaand aan de bijeenkomsten drie open vragen te stellen (voormeting) en hen na afloop dezelfde drie vragen opnieuw voor te leggen (nameting). Tussentijds, aan het einde van elke bijeenkomst, zijn de deelnemers gevraagd een aantal vragen te beantwoorden. Te zamen met de informatie over de deelnemers zelf levert een dergelijke opzet een beeld van:

- hoe de deelnemers veranderen in hun denken (vast te stellen onafhankelijk van de deelnemers zelf, door de onderzoeker)
- hoe de deelnemers zelf vinden dat zij veranderd zijn in hun denken (vast te stellen door de deelnemers zelf, onafhankelijk van de onderzoeker)
- hoe de deelnemers de sessies vonden
- de deelnemers zelf (leeftijd, afdeling, ervaring e.d.)

Echter om in staat te zijn uit de teksten af te leiden hoe de deelnemers veranderen in hun denken, dienen de teksten in zogenaamde cognitieve mappen ('cognitive maps') te worden omgevormd. Hoe een dergelijk constructie / transformatie heeft plaatsgevonden wordt hieronder, onder het kopje 'constructie van de variabelen' (zie ook hoofdstuk 6) beschreven.

*» constructie van de variabelen*

In deze studie trachten we inzicht te krijgen in (het verschuiven van) het denken van de deelnemers over slechtgestructureerde (beleids)problemen. Om de door de deelnemers geschreven teksten te kunnen analyseren, maken we gebruik van de 'cognitive mapping' benadering. Bij het omzetten van de teksten in cognitieve mappen dienen een aantal vaste transformatieregels (ook wel codeerregels genoemd) te worden uitgevoerd. Dit om te garanderen dat de omvorming op uniforme en consistente wijze plaatsvindt. De belangrijkste codeerregels zijn in bijlage 10 beschreven.

Naast het hanteren van duidelijke en eenduidige regels om uniformiteit en consistentie te waarborgen, dient tevens ex post te worden vastgesteld of toepassing van deze regels ook daadwerkelijk tot de gewenste uniformiteit en consistentie hebben geleid. Om dit te kunnen bepalen zijn in deze studie een deel van de teksten door drie codeurs gecodeerd (in cognitieve mappen omgevormd). De beschrijving van de opzet en uitkomsten van dit betrouwbaarheidsonderzoek worden in paragraaf 6.2.5 beschreven. Samenvattend kan gesteld worden dat de gemiddelde inter-codeur betrouwbaarheid met een score van 78 % acceptabel genoemd kan worden. De mate van overeenkomst tussen codeurs A en C is met 83 % zelfs zeer acceptabel. Hetzelfde geldt voor de intra-codeur betrouwbaarheid (de stabiliteit) van de onderzoeker. Deze bedraagt 82 %. Voldoende om wetenschappelijke conclusies op te baseren.

De op de vragenlijsten gebaseerde variabelen zijn tot stand gekomen op basis van a priori schaaltjes (factor analyse leidde niet tot acceptabele andersoortige oplossingen). Van te voren vastgestelde combinatie van vragen (a priori schaaltjes) hebben als basis gefungeerd voor de constructie van de variabelen en de aldus verkregen waarden voor deze variabelen. Met betrekking tot de variabelen die op de cognitieve mappen gebaseerd zijn (en dus niet op de vragenlijsten maar op de open vragen van de voor- en nameting) wordt een onderscheid tussen concepten en relaties gemaakt - de 'cognitive mapping' benadering is bij uitstek geschikt zijn om deze twee belangrijke componenten van ons denken over problemen analyseerbaar te maken. De aldus geconstrueerde variabelen kunnen als volgt worden samengevat:

<table>
<tr><td colspan="3">AFHANKELIJKE VARIABELEN</td></tr>
<tr><td>hoofdcriteria</td><td>gebaseerd op de cognitieve map</td><td>gebaseerd op de vragenlijsten</td></tr>
<tr><td rowspan="2">domein-specifiek (vragen II en IV)</td><td>endogeniteit (concepten en relaties)<br>exogeniteit (concepten en relaties)<br>externe model (concepten en relaties)</td><td>algemene domein-specifieke kennis<br>endogene en exogene kennis</td></tr>
<tr><td></td><td>inter-individuele verandering in domein-specifieke kennis</td></tr>
<tr><td>strategische (I en III)</td><td>feedbackloops<br>ketens<br>connectiviteit<br>time-phase relaties</td><td>feedbackloops<br>chaining<br>connectiviteit</td></tr>
<tr><td>bewustzijn van de positie van de ander (V)</td><td></td><td>bewustzijn van</td></tr>
</table>

<table>
<tr><td colspan="4">POTENTIEEL SPECIFICERENDE VARIABELEN</td></tr>
<tr><td colspan="4">achtergrondvariabelen</td></tr>
<tr><td>leeftijd<br>specialisme<br>tijdsinvestering</td><td>educationele achtergrond<br>jaren bij de organisatie werkzaam<br>aantal uren voorbereid</td><td>afdeling<br>beleidservaring<br>uren bijeenkomsten bijgewoond</td><td>organisatie<br>geslacht</td></tr>
<tr><td colspan="4">evaluatie van het programma</td></tr>
<tr><td>werkboek<br>subgroepsactiviteiten<br>plenaire sessie</td><td>zinvolheid<br>makkelijkheid<br>er van geleerd</td><td>interessantheid<br>tijd er aan besteed</td><td></td></tr>
</table>

## RESULTATEN

### • Algemeen

Het door de deelnemers in de voor- en nameting gebruikte aantal concepten en relaties daalde, zij het niet-significant. Het aantal concepten daalde van 24 naar 22, terwijl het aantal relaties tussen deze concepten van 26 naar 23 daalde. Dit geldt voor de on-treatment groep. Verder is opvallend dat maar liefst 50 % van de concepten en 78 % van de relaties die bij de nameting gebruikt worden nieuw zijn. Met andere woorden, de grootte van de cognitieve mappen verandert niet substantieel (de cognitieve mappen worden iets kleiner), maar de inhoud verandert wel en drastisch.

• Onderzoeksvragen
Ten aanzien van de eerste vier hoofdvragen kan op basis van de geschreven teksten en de daarop gebaseerde cognitieve mappen niet geconcludeerd worden dat de deelnemers substantieel meer strategische of domein-specifieke kennis krijgen. Ook wordt er door de deelnemers niet substantieel méér gebruik gemaakt van het ontwikkelde externe model. De inwisseling van concepten en relaties zoal hierboven beschreven, lijkt dus op enigzins 'willekeurige' wijze te gebeuren. Ook waar het gaat om de inter-individuele verandering in het denken over het probleem - de afname van de spreiding in het gebruik van concepten en relaties (inhoud), of feedbackloops, ketens, connectiviteit en tijdsindicatoren (vorm) - treedt er geen significante verschuiving op. Dit wordt ondersteund door de uitkomsten van de vragenlijsten: de gemiddelde scores op de 'kennis' variabelen zijn dusdanig dat niet tot toename kan worden geconcludeerd (zie hiervoor de figuren 7.3 en 7.5).

Met betrekking tot de vijfde onderzoeksvraag (het bewustzijn van de kennis en opvattingen van de andere deelnemers over het probleem) kan een gematigd positief antwoord worden gegeven. Op basis van de resultaten van de vragenlijst kan worden geconstateerd dat men van mening is meer inzicht in het denken van de anderen te hebben gekregen waar het gaat om het probleem van de kostenbeheersing in de gezondheidszorg.

• Specificerende variabelen
De bevindingen met betrekking tot de specificerende variabelen vallen uiteen in bevindigen die betrekking hebben op de waarde van de variabelen en de specificerende invloed van deze variabelen op de relatie tussen participatieve modelbouw en tot stand gebrachte effecten. De waarden van de variabelen worden in deze paragraaf beschreven. De specificerende invloed komt in de volgende paragraaf aan bod.

Met betrekking tot de opvattingen van de deelnemers over het programma, kan samenvattend worden geconstateerd dat de subgroepsactiviteiten en de plenaire onderdelen het predikaat goed hebben gekregen. De voorbereiding door middel van de werkboeken werd minder geslaagd genoemd, alhoewel er ook deelnemers geweest zijn die de werkboeken juist zeer zinvol en leerzaam hebben vonden. De participatieve modelbouw werd in het algemeen zinvol een interessant genoemd. Ook was men van mening er iets van geleerd te hebben (alhoewel niet al te veel). Over de lengte van het programma bestonden geen klachten.

Waar het gaat om de hoeveelheid tijd besteed aan het programma, kan vermeld worden dat deze in onze studie gemiddeld 13,5 uur bedragen heeft. Aanzienlijk minder dan de deelnemers van de Vennix studie, waar door studenten gemiddeld 22 uur en 55 minuten aan het programma besteed werd.

• Specificerende invloed van de specificerende variabelen
De invloed van bovenstaande specificerende variabelen op de relatie tussen participatieve modelbouw en opgetreden effecten kan als volgt worden samengevat (zie figuur 7.14 en bijlagen 12 en 13 voor een meer gedetailleerde beschrijving):

- mensen met een hogere opleiding veranderen meer (leren meer)
- zij die langer werkzaam zijn bij de organisatie, scoren beter ten aanzien van de domein-specifieke kennis verschuiving
- mensen die korter bij de organisatie werken scoren beter op de verschuiving in strategische kennis

- deelnemers van de afdeling Medische Dienst scoren beter bij de domein-specifieke kennis (en verschuiving)
- zij die minder tijd aan het programma besteden scoren beter bij de strategische kennis (verschuiving)
- oudere deelnemers verschuiven meer in het denken over het probleem
- bij de deelnemers die van mening waren van het programma geleerd te hebben, trad een verlaging van de strategische kennis scores op
- bij de deelnemers die van mening waren dat het programma zinvol was, werd een toename in de domein-specifieke kennis geconstateerd
- deelnemers die van mening waren dat het programma aan de korte kant was verhoogden hun domein-specifieke kennis, in tegenstelling tot de deelnemers die van mening waren dat het programma aan de lange kant was. Bij hen trad een verlaging van de domein-specifieke scores op

Deze hierboven samengevatte verschillen betreffen alleen de individuele verandering in het denken over het probleem van de kostenbeheersing in de gezondheidszorg. Dit omdat er ten aanzien van de inter-individuele verandering in het denken van de groep en de mate van bewust zijn van kennis van de andere deelnemers ten aanzien van het probleem géén specificerende effecten zijn gevonden.

• Vergelijking met de Vennix (1990) studie
Door een aantal van de variabelen van onze studie op een iet wat andere wijze te definiëren, zijn we in staat een deel van de resultaten van de Vennix te vergelijken met de resultaten van onze studie.

Opvallend resultaat van deze vergelijking is dat de in onze studie gevonden significante toename in 'precisie van concepten' (de proportie van modelconcepten) door de Vennix studie onderstreept wordt. In beide studies gaan de deelnemers door deelname beduidend meer concepten van het externe model gebruiken.

De vergelijking maakt ook zichtbaar dat in de Vennix studie een niet-slechtgestructureerd (economisch) probleem centraal heeft gestaan. De mate waarin van het externe (expert) model gebruik is gemaakt, is veel hoger (zelfs op de voormeting al) dan de mate waarin in de huidige studie op voor- en nameting van het externe model gebruik is gemaakt. Tevens is opvallend dat in de Vennix studie de mentale mappen een substantieel hogere dichtheid (totaal aantal relaties gedeeld door het totaal aantal concepten) hebben - er bestaan meer relaties tussen de concepten - de concepten zijn meer met elkaar verbonden (o.a. door verwijzingen).

Met betrekking tot de lengte van de ketens waarin men denkt, valt op dat experts (beleidsmakers) in langere ketens denken dan studenten (de Vennix studie). Zelfs in het geval studenten met redelijk goed gestructureerde problemen te maken hebben en beleidsmakers met slechtgestructureerde problemen geconfronteerd worden. Waar het gaat om de verschillen in de mate waarin de factor tijd (het aantal vertraagde relaties) door de deelnemers in het denken is opgenomen, is het waarschijnlijk dat de hogere scores in de Vennix studie het resultaat zijn van het expliciet aandacht besteden aan tijd en vertragingen. In de huidige studie zijn door de deelnemers geen analyses uitgevoerd of oefeningen gedaan die tot doel hadden het bestaan en de werking van vertragingen expliciet onder de aandacht te brengen.

## INTERPRETATIE VAN DE UITKOMSTEN

Bovenstaande uitkomsten geven aan dat de studie niet helemaal heeft opgeleverd wat er van verwacht werd. Hoewel er een model van de gezondheidszorg geconstrueerd is met behulp waarvan analyses kunnen worden uitgevoerd en inzicht kan worden verkregen in de complexiteit en dynamiek ervan (zie onder andere Post en Verburgh, 1993), zijn de deelnemers aan het programma niet in die mate in hun denken over het probleem veranderd dat van significante verschuivingen in het denken kan worden gesproken. Dit geldt voor zowel individuele als inter-individuele verschuivingen in het denken.

De vraag die op basis van deze uitkomsten gesteld dient te worden is de vraag naar het waarom van deze onverwachte uitkomsten. Daar onze studie zich niet alleen ten doel heeft gesteld de onderzoeksvragen te beantwoorden, doch ook voornemens was waar nodig de fundamenten van de studie kritisch te bekijken, valt de interpretatie uiteen in twee delen. In het eerste deel (kritische beschouwingen vanuit het zogenaamde rationalistische perspectief) worden de uitkomsten geïnterpreteerd binnen de gehanteerde aanpak. In het tweede deel (beschouwingen vanuit een non-trivial machine en actor perspectief) wordt de huidige benadering zelf ter discussie gesteld.

• Implicaties vanuit een *rationalistisch* perspectief

Bij een rationalistisch perspectief wordt gezocht naar wetmatigheden (zoals bijvoorbeeld de relatie tussen participatieve modelbouw en verandering in het denken) die onafhankelijk van de context (de actuele situatie) bestaan - dat wil zeggen, bevindingen die in alle gevallen waar zijn. Door intelligent gebruik te maken van een onderzoeksdesign kunnen context-invloeden worden uitgezuiverd zodat de 'objectieve', context-onafhankelijke wetmatigheden zichtbaar worden. Binnen het rationalistische perspectief wordt aan de onderzoeker met name een observator rol toegekend. De onderzoeker meet en aanschouwt het object (systeem) dat onderwerp van onderzoek is, om zo tot onderzoeker-onafhankelijke en context-onafhankelijke bevindingen te komen.

Vanuit dit rationalistische perspectief, dat min of meer het binnen de huidige studie gehanteerde perspectief is, kunnen bij de uitvoering van de huidige studie enkele kritische kanttekeningen worden gemaakt die een bijdrage kunnen leveren aan de verklaring van de wat (magere) resultaten. De te plaatsen kanttekeningen zijn:

- de evaluatie betreft slechts een verandering in het denken. Het daadwerkelijke handelen blijft hiermee buiten schot.
- de afwezigheid van een controlegroep maakt dat de effecten van de participatieve modelbouw niet onvoorwaardelijk aan de participatieve modelbouw kunnen worden toegeschreven.
- vergelijking met andere beleidsondersteunende methoden ontbreekt. De studie lijdt aan de door Bredemeyer en Greenblatt (1981) beschreven beperking: "So how's your wife Sam? Compared to what?"
- slechts de korte termijn effecten zijn in ogenschouw genomen daar de nameting slechts twee weken na de laatste bijeenkomst heeft plaatsgevonden. De vraag is of strategische kennis niet langer nodig heeft om geïncorporeerd te worden (Breuer and Kummer, 1990).
- de verschuiving in individueel en inter-individueel denken heeft plaatsgevonden op basis van dezelfde (formele) criteria. Misschien zou om veranderingen in het inter-individuele denken vast te stellen wel gebruik moeten worden gemaakt van groepsprodukten.
- door deelnemers te vragen twee keer dezelfde drie vragen te beantwoorden (voor- en nameting) zijn zij mogelijk minder gemotiveerd geweest bij het invullen van de nameting. Dit kan de gevolgen hebben gehad voor de hoogte van de scores op de nameting.
- de mate van participatie is beperkt geweest daar van een door de onderzoeker geconstrueerd voorlopig model gebruik is gemaakt én het aantal bijeenkomsten (tijdsinvestering) aan de magere kant was
- weinig is bekend over de mogelijkheid tot verandering van mentale mappen. Hierdoor kunnen de verschuivingen in het denken amper in enig perspectief geplaatst worden.
- het is de vraag of de mentale mappen die door middel van de 'cognitive mapping' benadering zijn geconstrueerd, in staat zijn om verfijnde verschuivingen in het denken vast te stellen. Zo zeggen de mentale mappen bijvoorbeeld niets over de status (zekerheid) van de door de deelnemers in de teksten uitgesproken conceptualisering.

- men kan zich afvragen of wel de juiste personen en organisatie aan de studie hebben deelgenomen. Door allerlei ontwikkelingen (fusie, plan Dekker, budgettering e.d.) is niet iedereen in staat geweest optimaal aan de sessies deel te nemen.
- een aantal van de deelnemers kunnen meer als 'executive decision-makers' dan als 'policy makers' (Vickers, 1965) beschouwd worden. Hierdoor hebben niet alle deelnemers in dezelfde mate gebruik gemaakt van de constructieve ruimte (de mogelijkheid om de 'rules of the game' te benoemen en te veranderen) die de participatieve modelbouw hen heeft geboden.

Op basis van bovenstaande beperkingen kan men en dient men **niet** te concluderen dat participatieve modelbouw **niet** in staat is om een verandering in het conceptualiseren van het probleem bij hen die er aan deelnemen tot stand te brengen. De beschreven limitaties geven echter wel aan dat er voor toekomstige onderzoeken ruimte voor verbetering bestaat - ruimte die zowel de plannings-, uitvoerings- als ook de evaluatiefase van het onderzoek betreft.

Dit neemt echter niet weg dat, gezien het feit dat geen van de eerste vier onderzoekshypothesen stand houdt, de vraag gesteld kan worden of de fundamenten van de aanpak niet ook in heroverweging dienen te worden genomen. Het ontbreken van (gevonden) effecten kan immers ook te wijten zijn aan een niet adequate conceptualisering en operationalisering van het wát van de invloed van participatieve modelbouw en de wijze waarop (het hóe) de evaluatie van de effecten heeft plaatsgevonden. In de volgende alinea zal een korte schets worden gegeven van een mogelijk alternatief theoretisch raamwerk met behulp waarvan het ontbreken van significante effecten kan worden verklaard en de relatie tussen participatieve modelbouw en conceptualisering kan worden uitgewerkt.

• Implicaties vanuit een *non-trivial machine en actor* perspectief

De in deze studie geïntroduceerde aanzet tot een mogelijke alternatief verklaringsmodel voor de afwezigheid van overtuigend empirisch bewijs voor de relatie tussen participatieve modelbouw en de conceptualisering van een beleidsprobleem onderscheidt zich van bovenstaand 'rationalistisch' perspectief door:

- de nadruk te leggen op de *context-afhankelijkheid* van de tijdens de sessies gegenereerde kennis. Door aan te geven dat bij beleidsondersteuning kennis binnen een handelingsperspectief plaats en betekenis krijgt, wordt het mogelijk om veranderingen in context als oorzaak voor een verschuiving in betekenis en gebruik van de gegenereerde kennis te zien.
- de notie van *proces-oriëntatie* te introduceren. Hierdoor worden de door participatieve modelbouw teweeg gebrachte effecten niet meer alleen geduid in termen van van te voren gespecificeerde en verwachte effecten verbonden aan een bepaald moment in de tijd, doch ook gezien als effecten die het resultaat zijn van een dynamisch proces waarbij, afhankelijk van de initiele 'waarden' van de deelnemers, verschillen in uitkomsten tot stand kunnen worden gebracht. Het is dan ook de vraag in hoeverre een dergelijk procesmatig karakter door middel van 'produkt'-evaluatie in kaart kan worden gebracht.
- het beheersingsdenken van de traditionele rationalistische benadering los te laten door in de relatie tussen participant en object van 'onderzoek' een *zelf-referentieel* karakter te onderkennen. Door het onderzoeksobject (het systeem van de gezondheidszorg) niet langer als een extern, te beheersen systeem, te beschouwen, doch te realiseren dat het deel uitmaken van het te bestuderen systeem gevolgen heeft voor het systeem, de bestudering, en degenen die de onderzoeksactiviteit uitvoeren, kan mogelijkerwijs inzichtelijk worden gemaakt dat waar het gaat om het wát van de evaluatie van participatieve modelbouw een verschuiving van *kennis* over het systeem in de richting van handelings- of reconstructieondersteunend vermogen op zijn plaats is om beleidsondersteuning en de bijdrage van participatieve modelbouw daaraan in kaart te kunnen brengen. Teneinde binnen een dergelijk zelf-referentiële relatie verbeterde

handelingsgeoriënteerde kennis van het systeem te kunnen krijgen, lijkt het wisselen van perspectief een veelbelovende oplossing. Door als betrokken actoren zowel een oberverende (outsider) als participerende (insider) perspectief te nemen, kan enerzijds de afstand tot het systeem worden genomen (outsider's perspectief) die nodig is om op andere wijze naar de processen die spelen te kunnen kijken, en anderzijds de betrokkenheid worden gecreëerd (insider's perspectief) zonder welke het niet mogelijk is om de betekenis van de processen te kunnen doorgronden. Het is in de wisseling van deze twee perspectieven dat betrokken actoren zich bewust kunnen worden van de relatie die zij hebben tot het systeem dat zij onderzoeken en waarbinnen ook zij opereren.

Het zelf-referentiële karakter van de relatie tussen object en subject (uitvoerder) van onderzoek betreft niet alleen de deelnemers aan de participatieve modelbouw in de context van beleidsondersteuning, doch mogelijkerwijs ook de relatie tussen onderzoek en onderzoeker. De noties en inzichten verkregen tijdens het onderzoek aangaande het object van onderzoek (het gebruik van participatieve modelbouw binnen de context van beleidsontwikkeling), hebben in deze studie ook invloed uitgeoefend op de manier waarop (beleids)onderzoek lijkt te moeten worden uitgevoerd - de manier waarop de onderzoeker zich tot objecten van onderzoek wenst te verhouden: wisselend in perspectief. Als zodanig heeft de kennis opgedaan tijdens de bestudering van de ontwikkeling en het gebruik van participatieve modelbouw gevolgen voor de kennis die ontwikkeld is en wordt in deze studie. De constatering dat in het uitoefenen van onderzoek men onderzoeker wordt krijgt hierdoor een diepere (zelf-referentiële) betekenis - zij zegt niet slechts iets over het object van onderzoek, doch ook over de onderzoeker zelf.

## GEVOLGEN VOOR HET ONTWERP VAN TOEKOMSTIGE PARTICIPATIEVE MODELBOUW BIJEENKOMSTEN

Vanuit een 'traditioneel' perspectief bezien, kunnen toekomstige toepassingen van participatieve modelbouw verbeterd worden door:

- de mate van participatie te vergroten. Hiertoe dient de hoeveelheid tijd te besteden aan het programma te worden vergroot
- een sterkere relatie tussen leerdoelen en angeboden detailprogramma tot stand te brengen. Met name de specifieke verschillen tussen strategische en domein-specifieke kenniselementen dienen in het aangeboden programma zichtbaar te worden gemaakt
- flexibel in te kunnen gaan op wensen van de deelnemers in het geval computer analyses onderdeel van het programma uitmaken. Vaststaande analyse-schema's dienen waar mogelijk vermeden te worden
- het gebruik van werkboeken ter voorbereiding op bijeenkomsten te vermijden
- vrijblijvendheid tijdens het programma uit te sluiten

Echter in het geval dat gekozen wordt voor een alternatief verklaringsschema voor de relatie tussen participatieve modelbouw en beleidsondersteuning (zoals uiteengezet in de beschrijving van het Non-Trivial Machine perspectief en de Actor benadering), lijkt het wenselijk dat toekomstige toepassingen van participatieve modelbouw de nadruk leggen op de mogelijkheid om wisselingen in perspectief mogelijk te maken. Hiertoe kunnen elementen van de wereld van gaming in het programma worden opgenomen (multi-actor simulatie)en/of kan de debriefing voor een dergelijke wisseling in perspectief gereserveerd worden. Het spreekt voor zich dat een dergelijke invulling van participatieve modelbouw als 'self-organizing learning environment' (Klabbers, 1990), niet zonder een zo groot mogelijke participatie van alle deelnemers kan daar tijdens de bijeenkomsten vorm en inhoud wordt gegeven ((re)constructie) aan het sociale systeem waarbinnen de betrokkenen participeren en gestalte aan het te voeren beleid dienen te geven.

## PRAKTISCHE GEVOLGEN

Een van de conclusies die uit deze studie kan worden getrokken is dat de evaluatie van de bijdrage die participatieve modelbouw aan beleidsontwikkeling levert nog een lange weg te gaan heeft. Niet alleen wat betreft de manier waarop toekomstige evaluaties onderzoekstechnisch dienen te worden uitgevoerd, doch ook wat betreft het theoretische wát van de bijdrage van participatieve modelbouw aan beleid(sontwikkeling). Slechts wanneer op bovenstaande vragen antwoord is gegeven kan nieuw empirisch onderzoek naar de effecten van participatieve modelbouw in overweging worden genomen.

# APPENDIX 1: QUESTIONNAIRE - SESSION 1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. Reading the workbook was | easy | | | | | | difficult |
| 2. The workbook was | interesting | | | | | | uninteresting |
| 3. From reading the workbook I have | learned a lot | | | | | | learned nothing at all |
| 4. Reading the workbook was | useful | | | | | | useless |
| 5. I consider the workbook to be | too long | | | | | | too short |
| 6. To prepare for today's session the workbook was | good | | | | | | bad |
| 7. The degree to which the content of workbook was in line with my daily work is | good | | | | | | bad |
| | | <1 | 1-2 | 2-3 | 3-4 | >4 | |
| 8. The amount of time I spent on reading the workbook was about (in hours) | | | | | | | |
| 9. The exercises in the workbook were | clear | | | | | | not clear |
| 10. The exercises in the workbook were | easy | | | | | | difficult |
| 11. The exercises in the workbook were | useful | | | | | | useless |
| 12. The assignments and discussions of the small-group session were | easy | | | | | | difficult |
| 13. The assignments and discussions of the small-group session were | interesting | | | | | | uninteresting |
| 14. From the assignments and discussions from the small-group session I have | learned a lot | | | | | | learned nothing at all |
| 15. The assignments and discussions of the small-group session were | useful | | | | | | useless |
| 16. The amount of time spent on the assignments and discussions of the small-group session was | too long | | | | | | too short |
| 17. The assignments and discussions of the small-group session have increased my knowledge of how my colleagues think about the health care system | agree | | | | | | disagree |
| 18. The plenary presentation and discussion was | interesting | | | | | | uninteresting |
| 19. From the plenary presentation and discussion I have | learned a lot | | | | | | learned nothing at all |
| 20. The plenary presentation and discussion was | useful | | | | | | useless |
| 21. The amount of time spent on the plenary presentation and discussion was | too long | | | | | | too short |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 22. Altogether, the components of the program were in line with each other | agree | ☐ | ☐ | ☐ | ☐ | ☐ | disagree |
| 23. Altogether, the session has resulted in an increase in my knowledge about the health care system | agree | ☐ | ☐ | ☐ | ☐ | ☐ | disagree |
| 24. Altogether, I have acquired more knowledge of the ideas and opinions of the other participants by taking part in this session | agree | ☐ | ☐ | ☐ | ☐ | ☐ | disagree |
| 25. This session has succeeded in making me exchange my thoughts intensively with the other participants | agree | ☐ | ☐ | ☐ | ☐ | ☐ | disagree |

# APPENDIX 2: QUESTIONNAIRE - SESSION 2

| | | | |
|---|---|---|---|
| 1. Reading the workbook was | easy | ☐☐☐☐☐ | difficult |
| 2. The workbook was | interesting | ☐☐☐☐☐ | uninteresting |
| 3. From reading the workbook I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 4. Reading the workbook was | useful | ☐☐☐☐☐ | useless |
| 5. I consider the workbook to be | too long | ☐☐☐☐☐ | too short |
| 6. To prepare for today's session the workbook was | good | ☐☐☐☐☐ | bad |
| 7. The degree to which the content of workbook was in line with my daily work is | good | ☐☐☐☐☐ | bad |
| 8. The amount of time I spent on reading the workbook was about (in hours) | | <1 1-1.5 1.5-2 2-2.5 >2.5<br>☐☐☐☐☐ | |
| 9. The exercises in the workbook were | clear | ☐☐☐☐☐ | not clear |
| 10. The exercises in the workbook were | easy | ☐☐☐☐☐ | difficult |
| 11. The exercises in the workbook were | useful | ☐☐☐☐☐ | useless |
| 12. The assignments and discussions of the small-group session were | easy | ☐☐☐☐☐ | difficult |
| 13. The assignments and discussions of the small-group session were | interesting | ☐☐☐☐☐ | uninteresting |
| 14. From the assignments and discussions from the small-group session I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 15. The assignments and discussions of the small-group session were | useful | ☐☐☐☐☐ | useless |
| 16. The amount of time spent on the assignments and discussions of the small-group session was | too long | ☐☐☐☐☐ | too short |
| 17. The assignments and discussions of the small-group session have increased my knowledge of how my colleagues think about the health care system | agree | ☐☐☐☐☐ | disagree |
| 18. The plenary presentation and discussion was | interesting | ☐☐☐☐☐ | uninteresting |
| 19. From the plenary presentation and discussion I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 20. The plenary presentation and discussion was | useful | ☐☐☐☐☐ | useless |
| 21. The amount of time spent on the plenary presentation and discussion was | too long | ☐☐☐☐☐ | too short |

| | | | |
|---|---|---|---|
| 22. Altogether, the components of the program were in line with each other | agree | ☐☐☐☐☐ | disagree |
| 23. Altogether, the session has resulted in an increase in my knowledge about the health care system | agree | ☐☐☐☐☐ | disagree |
| 24. Altogether, I have acquired more knowledge of the ideas and opinions of the other participants by taking part in this session | agree | ☐☐☐☐☐ | disagree |
| 25. This session has succeeded in making me exchange my thoughts intensively with the other participants | agree | ☐☐☐☐☐ | disagree |

26. Age
0 20-29
0 30-39
0 40-49
0 50-59
0 60 and older

27. What is the highest education you've taken?
0 secondary education
0 intermediate vocational education - in the field of: ....................
0 higher vocational education - in the field of: ................
0 university - in the field of:...................

28. How long have you been with this organization?
0 0-1 years
0 2-3 years
0 4-5 years
0 6-10 years
0 > 10 years

29. The component of the health care system that you know most of is:
0 the general practitioner component
0 the medical specialist component
0 the hospital component

30. How much of the work that you do is concerned with policy making?
0 > 80 % of the work that I do
0 60-80 % of the work that I do
0 40-59 % of the work that I do
0 20-39 % of the work that I do
0 < 20 % of the work that I do

# APPENDIX 3: QUESTIONNAIRE - SESSION 3

| | | | |
|---|---|---|---|
| 1. Reading the workbook was | easy | ☐☐☐☐☐ | difficult |
| 2. The workbook was | interesting | ☐☐☐☐☐ | uninteresting |
| 3. From reading the workbook I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 4. Reading the workbook was | useful | ☐☐☐☐☐ | useless |
| 5. I consider the workbook to be | too long | ☐☐☐☐☐ | too short |
| 6. To prepare for today's session the workbook was | good | ☐☐☐☐☐ | bad |
| 7. The degree to which the content of workbook was in line with my daily work is | good | ☐☐☐☐☐ | bad |
| 8. The amount of time I spent on reading the workbook was about (in hours) | | <1 1-1.5 1.5-2 2-2.5 >2.5<br>☐☐☐☐☐ | |
| 9. The exercises in the workbook were | clear | ☐☐☐☐☐ | not clear |
| 10. The exercises in the workbook were | easy | ☐☐☐☐☐ | difficult |
| 11. The exercises in the workbook were | useful | ☐☐☐☐☐ | useless |
| 12. The assignments and discussions of the small-group session were | easy | ☐☐☐☐☐ | difficult |
| 13. The assignments and discussions of the small-group session were | interesting | ☐☐☐☐☐ | uninteresting |
| 14. From the assignments and discussions from the small-group session I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 15. The assignments and discussions of the small-group session were | useful | ☐☐☐☐☐ | useless |
| 16. The amount of time spent on the assignments and discussions of the small-group session was | too long | ☐☐☐☐☐ | too short |
| 17. The assignments and discussions of the small-group session have increased my knowledge of how my colleagues think about the health care system | agree | ☐☐☐☐☐ | disagree |
| 18. The plenary presentation and discussion was | interesting | ☐☐☐☐☐ | uninteresting |
| 19. From the plenary presentation and discussion I have | learned a lot | ☐☐☐☐☐ | learned nothing at all |
| 20. The plenary presentation and discussion was | useful | ☐☐☐☐☐ | useless |
| 21. The amount of time spent on the plenary presentation and discussion was | too long | ☐☐☐☐☐ | too short |

22. Altogether, the components of the program were in line with each other — agree ☐☐☐☐☐ disagree

23. Altogether, the session has resulted in an increase in my knowledge about the health care system — agree ☐☐☐☐☐ disagree

24. Altogether, I have acquired more knowledge of the ideas and opinions of the other participants by taking part in this session — agree ☐☐☐☐☐ disagree

25. This session has succeeded in making me exchange my thoughts intensively with the other participants — agree ☐☐☐☐☐ disagree

26. The participative policy modelling sessions have made me see (different) connections between the various components of the health care system — agree ☐☐☐☐☐ disagree

27. Due to the participative policy modelling sessions, I have acquired a deeper understanding of the causes of potential developments and the effects of potentially policy measures taken in the health care system — agree ☐☐☐☐☐ disagree

28. The participative policy modelling sessions have resulted in an increase in the knowledge I have of aspects of the health care system that fall beyond my daily work-activities — agree ☐☐☐☐☐ disagre

29. The participative policy modelling sessions have made me feel that feedbackprocesses are important in considering a complex system such as the health care system — agree ☐☐☐☐☐ disagree

30. The participative policy modelling sessions enabled me to state more precisely what is and is not important for the problem of cost containment in the health care system — agree ☐☐☐☐☐ disagree

31. By communicating to people with a somewhat different view on the health care system, I have acquired new ideas and insight — agree ☐☐☐☐☐ disagree

32. The participative policy modelling sessions have resulted in a broader perspective on the health care system — agree ☐☐☐☐☐ disagree

33. Due to the participative policy modelling sessions, I have acquired an idea of how my colleagues think about the health care system — agree ☐☐☐☐☐ disagree

34. Due to the participative policy modelling sessions, I know about what I agree and disagree with my colleagues — agree ☐☐☐☐☐ disagre

35. It is my impression that due to the participative policy modelling sessions we have become more alike in the way in which we think about the health care system — agree ☐☐☐☐☐ disagree

# APPENDIX 4: QUESTIONNAIRE - SESSION 4

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 36. The fourth session was | easy | | | | | | difficult |
| 37. The fourth session was | interesting | | | | | | uninteresting |
| 38. From the fourth session I have | learned at lot | | | | | | learned nothing at all |
| 39. The fourth session was | useful | | | | | | useless |
| 40. The components of the program of the fourth session were related to each other in a meaningful way | agree | | | | | | disagree |
| 41. The fourth session has increased my knowledge | agree | | | | | | disagree |
| 42. Due to the participative policy modelling sessions, I have acquired an idea of how my colleagues think about the health care system | agree | | | | | | disagree |
| 43. This session has succeeded in making me exchange my thoughts intensively with the other participants | agree | | | | | | disagree |

# APPENDIX 5: PRE- AND POSTTEST

In het licht van de verwachte groei van de kosten van de gezondheidszorg zijn reeds een groot aantal mogelijke beleidsmaatregelen naar voren gebracht. We willen u, op grond van uw ervaring in en kennis van het gezondheidszorgsysteem, vragen om 'a titre personel' met betrekking tot drie van deze beleidsmaatregelen aan te geven welke effecten u verwacht dat de desbetreffende maatregel voor de gezondheidszorg zal hebben. Mocht u meer ruimte nodig hebben dan de in de blokken gereserveerde ruimte, dan kunt u aan de achterkant van het blad verder schrijven.

De door u gegeven antwoorden zullen gebruikt worden in het aan het seminar 'Patientenstromen in de gezondheidszorg' (waar de afdelingen 'Overeenkomsten & Tarieven', 'Medische Dienst' en 'Informatiebeheer' te zamen aan deelnemen) gekoppelde onderzoek. Het valt onder de op uw schema van activiteiten vermelde 'introductie'. Ik verwacht dan ook dat u, zoals op dat schema is aangegeven, ongeveer een half uur nodig zult hebben om de vragen te beantwoorden. Voor alle duidelijkheid, met de antwoorden zal vertrouwelijk worden omgesprongen Dat wij u toch willen vragen om rechtsboven uw naam in te vullen heeft te maken met de verwerking van de antwoorden.

Op dinsdag 24 september kom ik naar [ ] om de antwoorden op de vragen op te halen en u het werkboek, dat u ter voorbereiding op de bijeenkomst dient door te nemen, te overhandigen. Mocht u er op dinsdag 25 september niet zijn, dan wil ik u vriendelijk verzoeken om voor die datum de antwoorden bij het secretariaat van de afdeling 'Medische Dienst' of de afdeling 'Overeenkomsten & Tarieven' in te leveren. Bij voorbaat mijn dank en in ieder geval tot de eerste bijeenkomst.

Luc Verburgh.

Vragen:

1. Een eerste mogelijke maatregel ter beheersing van de kosten van de gezondheidszorg bestaat uit het *vervangen van het abonnementstarief voor ziekenfondspatienten door een financieringsstelsel waarin een combinatie van een abonnementstarief en een verrichtingentarief gehanteerd wordt*. Stel dat deze maatregel wordt ingevoerd en dat 10 procent van het abonnementstarief vervangen wordt door een verrichtingentarief, wat denkt u dat de gevolgen voor de gezondheidszorg zullen zijn? Leg in uw antwoord ook de nadruk op waarom u denkt dat deze effecten zullen optreden.

2. Een tweede mogelijke maatregel ter beheersing van de verwachte groei van de kosten van de gezondheidszorg bestaat uit het *aanstellen van specialisten in loondienst* Stel dat deze maatregel wordt ingevoerd (zie af van de politieke haalbaarheid ervan), wat denkt u dat dit voor gevolgen voor de gezondheidszorg zal hebben? Leg in uw antwoord de nadruk op waarom u denkt dat deze effecten zullen optreden.

3. Een derde mogelijke maatregel ter beheersing van de kosten bestaat uit het *reduceren van de gemiddelde ligduur (verpleegduur)* Stel dat door protocollering de gemiddelde ligduur met 10 procent kan dalen, wat denkt u dat de gevolgen voor de gezondheidszorg zullen zijn? Leg in uw antwoord ook de nadruk op waarom u denkt dat deze effecten zullen optreden.

# APPENDIX 5: PRE- AND POSTTEST (english version)

In light of the expected growth of the costs of health care, a number of possible measures have been put forward and discussed. We would like to ask you, based on your knowledge and experience of the health care system, to describe what you think the effects of three of these measures on the health care system will be. We would like you to answer these questions based on your personal view on what the effects will be rather than based on what may be considered as an ' official point of view'. In case you need more space to write down your answer than the boxes we reserved for you, don't hesitate to write on the back of the page.

The answers given by you will be used for research connected to the seminar 'Patients flows in the health care system' (in which the Financial-legal department and Medical department will take part together). It is being referred to as ' Introduction' in the time-table that was sent to you earlier on. It is our expectation that about a half an hour is needed to complete the questions. To avoid any misunderstanding, the answers will be dealt with confidentially. The reason why we ask you to write down your name at the top right hand corner, is to be able to use your answer in the research.

On tuesday the 24th of September, I'll visit [ ] to collect your answers and to hand out the workbook that needs to be read in preparation of the first session. In case you are not in at the 24th, I would like you to give your answers to the secretary of the Financial-legal department or the Medical department. Thanking you in anticipation and looking forward to the first session,

Yours sincerely,

Luc Verburgh.

Questions:

1. A first possible measure to contain the costs of health care is to *replace the fixed annual fee for ' ziekenfonds' patients by a fee which is a combination of a fixed fee and a fee per medical transaction*. Suppose this measure is introduced and 10 per cent of the fixed fee is substituted by a fee per transaction, what do you think the effects for the health care system will be? Focus in your answer also on the reason why you think the effects will be as you describe them.

2. A second possible measure to contain the costs of health care is to have *medical specialist work for hospitals rather than have their own practice in hospitals* (i.e. give them a fixed annual income). Suppose this measure is being introduced (ignore the issue whether this measure is politically attainable), what do you think the effects for the health care system will be? Focus in your answer also on the reason why you think the effects will be as you describe them.

3. A third possible measure to contain the costs exists of reducing the *average length of stay in hospital*. Suppose that introducing protocols will result in a 10 per cent decresase of the average length of stay, what do you think the effects for the health care system will be? Focus in your answer also on the reason why you think the effects will be as you describe them.

# APPENDIX 6: THE MODEL OF THE DUTCH HEALTH CARE SYSTEM (IThink model used in Chapter 5)

patho(t) = patho(t - dt) + (disgpho + dismsho + dishosho - conhogp) * dt
INIT patho = 200347

patgp(t) = patgp(t - dt) + (conhogp + conwgpgp - refgpwms - ordgpwgp - disgpho) * dt
INIT patgp = 13321

pathwgp(t) = pathwgp(t - dt) + (ordgpwgp + refhoswgp + refmswgp - conwgpgp) * dt
INIT pathwgp = 14122

patwms(t) = patwms(t - dt) + (refgpwms + ordmswms + refhoswms - conwsms) * dt
INIT patwms = 40147

patms(t) = patms(t - dt) + (conwsms - refmswhos - ordmswms - dismsho - refmswgp) * dt
INIT patms = 8920

patwhos(t) = patwhos(t - dt) + (refmswhos - admwhoshos) * dt
INIT patwhos = 6428

pathos(t) = pathos(t - dt) + (admwhoshos - refhoswms - refhoswgp - dishosho) * dt
INIT pathos = 1304

pcconhogp = .04689
conhogp = pcconhogp*patho
gps = 102
workloadgp = patgp/gps
pcrefgpwms = GRAPH(workloadgp)
(110, 0.124), (114, 0.124), (118, 0.125), (122, 0.125), (126, 0.127), (130, 0.130), (134, 0.132), (138, 0.136), (142, 0.138), (146, 0.138), (150, 0.138)
refgpwms
ms = 171
workloadms = patms/ms
atwms = GRAPH(workloadms)
(40.0, 4.29), (42.0, 4.29), (44.0, 4.30), (46.0, 4.36), (48.0, 4.43), (50.0, 4.47), (52.0, 4.50), (54.0, 4.51), (56.0, 4.53), (58.0, 4.58), (60.0, 4.65)
conwsms = patwms/atwms
effect1 = GRAPH(workloadms)
(40.0, 0.0879), (42.0, 0.0880), (44.0, 0.0881), (46.0, 0.0885), (48.0, 0.0893), (50.0, 0.0897), (52.0, 0.0900), (54.0, 0.0910), (56.0, 0.0930), (58.0, 0.0933), (60.0, 0.0935)
pcrefmswhos = effect1
refmswhos = pcrefmswhos*patms
beds = 1635
occupation = pathos/beds
atwhos = GRAPH(occupation)
(0.700, 7.60), (0.725, 7.62), (0.750, 7.89), (0.775, 8.00), (0.800, 8.00), (0.825, 8.00), (0.850, 8.00), (0.875, 8.06), (0.900, 8.12), (0.925, 8.17), (0.950, 8.56)
admwhoshos = patwhos/atwhos
pcordgpwgp = GRAPH(workloadgp)
(110, 0.290), (114, 0.289), (118, 0.289), (122, 0.287), (126, 0.284), (130, 0.280), (134, 0.274), (138, 0.265), (142, 0.261), (146, 0.260), (150, 0.260)
ordgpwgp = pcordgpwgp*patgp

atlwgp = GRAPH(workloadgp)
(110, 3.30), (114, 3.33), (118, 3.36), (122, 3.48), (126, 3.56), (130, 3.59), (134, 3.63), (138, 3.71), (142, 3.86), (146, 3.93), (150, 4.00)
conwgpgp = pathwgp/atlwgp
pcordmswms = GRAPH(workloadms)
(40.0, 0.756), (42.0, 0.754), (44.0, 0.752), (46.0, 0.748), (48.0, 0.744), (50.0, 0.742), (52.0, 0.725), (54.0, 0.719), (56.0, 0.709), (58.0, 0.704), (60.0, 0.700)
ordmswms = pcordmswms*patms
pcrefhoswms = GRAPH(workloadms)
(40.0, 0.933), (42.0, 0.932), (44.0, 0.930), (46.0, 0.925), (48.0, 0.919), (50.0, 0.911), (52.0, 0.900), (54.0, 0.889), (56.0, 0.882), (58.0, 0.876), (60.0, 0.872)
l_of_stay = 11.36
weeks = l_of_stay/7
refhoswms = pcrefhoswms*(pathos/weeks)
pcrefhoswgp = .025
refhoswgp = pcrefhoswgp*(pathos/weeks)
pcdisgpho = 1-(pcordgpwgp+pcrefgpwms)
disgpho = pcdisgpho*patgp
pcrefmswgp = GRAPH(workloadms)
(40.0, 0.0187), (42.0, 0.0187), (44.0, 0.0189), (46.0, 0.0191), (48.0, 0.0195), (50.0, 0.0198), (52.0, 0.0200), (54.0, 0.0209), (56.0, 0.0220), (58.0, 0.0240), (60.0, 0.0260)
pcdisms = 1-(pcordmswms+pcrefmswhos+pcrefmswgp)
dismsho = pcdisms*patms
pcdishosho = 1-(pcrefhoswms+.025)
dishosho = pcdishosho*(pathos/weeks)
refmswgp = pcrefmswgp*patms

# APPENDIX 7: THE MODEL OF THE DUTCH HEALTH CARE SYSTEM (DYNAMO model used in the sessions)

## List of abbreviation used (in alphabetical order)

| | | |
|---|---|---|
| AANT_BD | = | number of hospital beds |
| AANT_ HA | = | number of general practitioners |
| AANT_SP | = | number of medical specialists |
| AKVRZKPC | = | number of clinical treatments per contact |
| APVRSPPC | = | number of poly-clinical treatments per contact |
| ATLWH | = | actual waiting time for general practitioner |
| ATLWS | = | actual waiting time for medical specialist |
| ATLWZ | = | actual waiting time for being admitted into hospital |
| AVRSHAPC | = | number of prescriptions per general practitioner per contact |
| AVRSSPC | = | number of prescriptions per medical specialist per contact |
| BZGRWZK | = | actual rate of occupancy |
| BZGREZK | = | perceived rate of occupancy |
| CONTHHA | = | consulting a general practitioner |
| CONWSSP | = | consulting a medical specialist |
| DRUKEHA | = | perceived workload general practitioner |
| DRUKWHA | = | actual worload general practitioner |
| DRUKESP | = | perceived workload medical specialist |
| DRUKWSP | = | actual workload medical specialist |
| DVWHAWS | = | referred by the general practitioner |
| DVWSPWZ | = | referred by the medical specialist |
| GROEI | = | growth of population |
| GTLWH | = | average waiting time for general practitioner |
| GTLWS | = | average waiting time for medical specialist |
| GTLWZ | = | average waiting time for being admitted into hospital |
| KOABOHA | = | costs of fixed fee for general practitioner per 'ziekenfonds' patient |
| KOCONHA | = | costs of privately insured patients consulting a general practitioner |
| KOHHKSP | = | costs of patients ordered back by medical specialists |
| KOKVRZK | = | costs of clinical treeatment |
| KOLIGZK | = | costs of being nursed at a hospital |
| KOPVRSP | = | costs of poly-clinical treatment |
| KOTOTAL | = | total costs of health care |
| KOVRSHA | = | costs of prescriptions of general practitioners |
| KOVRSSP | = | costs pf prescriptions of medical specialists |
| KOVWKSP | = | costs of patients referred to medical specialist |
| LIGDRZK | = | influence of rate of occupancy on length of stay in hospital |
| LIGDUUR | = | length of stay in hospital |
| ONTHATH | = | dismissed by general practitioner |
| ONTSPTH | = | dismissed by medical specialist |
| ONTZKTH | = | dismissed from hospital |
| OPNWZZK | = | admitted into hospital |
| PATN_HA | = | number of patients consulting a general practitioner |

| | | |
|---|---|---|
| PATN_SP | = | number of patients consulting a medical specialist |
| PATN_TH | = | number of patients at home |
| PATN_WH | = | number of patients waiting for another treatment by the general practitioner |
| PATN_WS | = | number of patients waiting for another treatment by the medical specialist |
| PATN_ZK | = | number of patients in hospital |
| | | |
| PCCONTH | = | consultation percentage |
| PCDVWHA | = | referral percentage general practitioner |
| PCDVWSP | = | referral percentage medical specialist |
| PCONTHA | = | dismissal percentage general practitioner |
| PCONTSP | = | dismissal percentage medical specialist |
| PCONTZK | = | dismissal percentage hospital |
| PCTBSHA | = | order bacl percentage general practitioner |
| PCTBSSP | = | order back percentage medical specialist |
| PCTVWSP | = | referred back percentage medical specialist to general practitioner |
| PCTVWZK | = | referred back percentage hospital to general practitioner |
| PCTVWZS | = | referred back percentage hospital to medical specialist |
| | | |
| PRABOHA | = | fixed fee for 'ziekenfonds' patients |
| PRCONHA | = | price of a private consultation of the general practitioner |
| PRHHKSP | = | price of being ordered back for another consultation of the medical specialist |
| PRLIGZK | = | price of a day in hospital |
| PRKVRSP | = | average price of average clinical treatment |
| PRVRSHA | = | average price of average poly-clinical treatment |
| PRVRSSP | = | average price of average prescription by the medical specialist |
| PRVWKSP | = | price of being referred to the medical specialist |
| | | |
| TBSHAWH | = | ordered back by the general practitioner |
| TBSSPWS | = | ordered back by the medical specialist |
| TVWSPWH | = | referred back from the medical specialist to the general practitioner |
| TVWZKWS | = | referred back from hospital to the medical specialist |
| TVWZKWH | = | referred back from hospital to the general practitioner |
| | | |
| VLABOHA | = | number of patients that are insured at a regional health care insurance organization |
| VLCONHA | = | number of privately insured patients |
| VLLIGZK | = | number of beddays (in hospital) |
| VLKVRZK | = | number of clinical treatments by medical specialists |
| VLPVRSP | = | number of poly-clinical treatments by medical specialists |
| VLVRSHA | = | number of prescriptions by general practitioners |
| VLVRSSP | = | number of prescriptions by medical specialists |

## PATIENTS NOT UNDER TREATMENT

```
L  PATN_TH.K     =  PATN_TH.J+DT*(GROEI.JK+ONTHATH.JK+ONTSPTH.JK+
                    ONTZKTHJK-CONTHHAJK)
N  PATN_TH       =  200200
R  CONTHHA.KL    =  PCCONTH.K*PATN_TH.K
A  PCCONTH.K     =  .04696+HACONDR.K
```

## PATIENTS TREATED BY THE GENERAL PRACTITIONER

```
L  PATN_HA.K     =  PATN_HAJ+DT*(CONTHHAJK+CONWHHAJK-ONTHATHJK-
                    TBSHAWHJK-DVWHAWSJK)
N  PATN_HA       =  13340
```

```
R  ONTHATH.KL    =  PCONTHA.K*PATN_HA.K
N  PCONTHA       =  .59
R  TBSHAWH.KL    =  PCTBSHA.K*PATN_HA.K
N  PCTBSHA       =  .28
A  PCTBSHA.K     =  .28+HATBSDR.K
A  HATBSDR.K     =  TABHL (PCDRHA2,DRUKEHA.K,110,150,5)
T  PCDRHA2       =  .050/.025/.020/.005/.001/-.002/-.010/-.025/-.050
A  PCONTHA.K     =  1-(PCTBSHA.K+PCDVWHA.K)
R  DVWHAWS.KL    =  PCDVWHA.K*PATN_HA.K
N  PCDVWHA       =  .13
A  PCDVWHA.K     =  .13+HADVWDR.K
A  HADVWDR.K     =  TABHL(PCDRHA1,DRUKEHA,110,155,5)
T  PCDRHA1       =  -.030/-.015/-.008/-.002/-.001/.002/.006/.015/.025/.040
```

## PATIENTS WAITING FOR TREATMENT BY THE GENERAL PRACTITIONER

```
L  PATN_WH.K     =  PATN_WHJ+DT*(TBSHAWH.JK+TVWZKWHJK+TVWSPWHJK-
                    CONWHHAJK)
N  PATN_WH       =  14200
R  CONWHHA.KL    =  PATN_WH.K/GTLWH.K
N  GTLWH         =  3.6
A  ATLWH.K       =  3.6+HAATLWHDR.K
A  HAATLWHDR.K   =  TABHL(PCDRHA4,DRUKEHA.K,110,155,5)
T  PCDRHA4       =  -.30/-.15/-.08/-.02/-.005/.03/.07/.15/.25/.40
N  ATLWH         =  3.6
A  GTLWH.K       =  DELAY1(ATLWH.K,3.6)
```

## PATIENTS WAITING FOR TREATMENT BY THE MEDICAL SPECIALIST

```
L  PATN_WS.K     =  PATN_WS.J+DT*(TBSSPWS.JK+TVWZKWS.JK+DVWHAWS.JK-
                    CONWSSP.JK)
N  PATN_WS       =  40200
R  CONWSSP.KL    =  PATN_WS.K/GTLWS.K
N  GTLWS         =  4.5
A  ATLWS.K       =  4.5+SPATLWSDR.K
A  SPATLSDR.K    =  TABHL(PCDRSP5,DRUKESP.K,40,60,2.5)
T  PCDRSP5       =  -2.0/-1.9/-1.5/-.7/-.4/.00//.4/.7/1.5
N  ATLWS         =  4.5
A  GTLWS.K       =  DELAY1(ATLWS.K,4.5)
```

## PATIENTS TREATED BY THE MEDICAL SPECIALIST

```
L  PATN_SP.K     =  PATN_SP.J+DT*(CONWSSP.JK-TBSSPWS.JK-ONTSPTH.JK-
                    DVWSPWZ.JK-TVWSPWH.JK)
N  PATN_SP       =  8937
R  TBSSPWS.KL    =  PCTBSSP.K*PATN_SP.K
N  PCTBSSP       =  .7974
A  PCTBSSP       =  .7974+SPTBSDR.K
A  SPTBSDR.K     =  TABHL(PCDRSP1,DRUKESP.K,40,60,2.5)
T  PCDRSP1       =  .060/.030/.016/.014/.01/-.000/-.01/-.014/-.030
R  ONTSPTH.KL    =  PCONTSP.K*PATN_SP.K
A  PCONTSP.K     =  1.00-(PCTBSSP.K+PCDVWSP.K+PCTVWSP.K)
R  DVWSPWZ.KL    =  PCDVWSP.K*PATN_SP.K
N  PCDVWSP       =  .09
A  PCDVWSP.K     =  .09+SPDVWDR.K+ZKDVWDR.K
```

```
S   SPDVWDR.K     =  TABHL(PCDRSP2,DRUKESP.K,40,60,2.5)
T   PCDRSP2       =  -.013/-.009/-.006/-.003/-.001/.000/.001/.003/.006
A   ZKDVWDR.K     =  TABHL(PCDRZK1,BZGREZK.K,.70,.95,.025)
T   PCDRZK1       =  .002/.0015/.0005/-.005/-.001/-.0015/-.002/-.003/-.005/-.01
R   TVWSPWH.KL    =  PCTVWSP.K*(PATN_SP.K/1)
N   PCTVWSP       =  .02
A   PCTVWSP.K     =  .02+SPTVWDR.K
A   SPTVWDR.K     =  TABHL(PCDRSP6,DRUKESP.K,40,60,2.5)
T   PCDRSP6       =  -.0013/-.0009/-.0006/-.0003/-.0001/.000/.0001/.0003/.0006
```

**PATIENTS WAITING FOR ADMISSION INTO HOSPITAL**

```
L   PATN_WZ.K     =  PATN_WZ.J+DT*(DVWSPWZ.JK-OPNWZZK.JK)
N   PATN_WZ       =  6426
R   OPNWZZK.KL    =  PATN_WZ.K/GTLWZ.K
N   GTLWZ         =  8
A   GTLWZ.K       =  DELAY1(ATLWZ.K,8)
A   ATLWZ.K       =  8+ZKGTLWZ.K
A   ZKGTLWZ.K     =  TABHL(PCDRZK2,BZGREZK.K,.70,.90,.025)
T   PCDRZK2       =  -.4/-.3/-.2/-.05/.05/.1/.2/.5/.75/1.1/1.75
```

**PATIENTS ADMITTED INTO HOSPITAL**

```
L   PATN_ZK.K     =  PATN_ZK.J+DT*(OPNWZZK.JK-TVWZKWS.JK-TVWZKWH.JK-
                     ONTZKTH.JK)
N   PATN_ZK       =  1310
R   TVWZKWS.KL    =  PCTVWZS.K*(PATN_ZK.K/LIGDUUR.K)
N   PCTVWZS       =  .10
A   PCTVWZS.K     =  TABHL(A1,DRUKESP.K,40,60,2.5)
T   A1            =  .09/.095/.097/.099/.0997/.101/.103/.105/.11
R   TVWZKWH.KL    =  .025*(PATN_ZK/LIGDUUR.K)
C   PCTVWZH       =  .025
R   ONTZKTH.KL    =  PCONTZK.K*(PATN_ZK.K/LIGDUUR.K)
A   PCONTZK.K     =  1-(PCTVWZS.K+.025)
A   LIGDUUR.K     =  ((11.36+LIGDRZK.K)/7)
A   LIGDRZK.K     =  TABHL(PCDRZK3,BZGREZK.K,.70,.90,.025)
T   PCDRZK3       =  1.2/.7/.3/.05/-.05/-.3/-.5/-.7/-1/-1.4/-1.9/-2.8
```

**WORKLOAD AND RATE OF OCCUPANCY**

WORKLOAD GENERAL PRACTITIONER

```
A   DRUKWHA.K     =  PATN_HA.K/AANT_HA
A   DRUKEHA.K     =  SMOOTH(DRUKWHA.K,4)
```

WORKLOAD MEDICAL SPECIALIST

```
A   DRUKWSP.K     =  PATN_SP.K/AANT_SP
A   DRUKESP.K     =  SMOOTH(DRUKWSP.K,4)
```

RATE OF OCCUPANCY

```
A   BZGRWZK.K     =  PATN_ZK.K/AANT_BD
A   BZGREZK.K     =  SMOOTH(BZGRWZK.K,1)
```

**VOLUME OF MEDICAL PRODUCTION**

CONSULTATION OF GENERAL PRACTITIONER

| | | | |
|---|---|---|---|
| A | VLABOHA.K | = | (1-.3247)*PATN_TO.K |
| A | VLCONHA.K | = | .3247*PATN_HA.K |

PRESCRIPTIONS GENERAL PRACTITIONER

| | | | |
|---|---|---|---|
| A | VLVRSHA.K | = | PATN_HA.K*AVRSHAPC.K |
| N | AVRSHAPC | = | 2.30 |
| A | AVRSHAPC.K | = | 2.30+HAVRSDR.K |
| A | HAVRSDR.K | = | TABHL(A4,DRUKEHA.K,110,155,5) |
| T | A4 | = | -.03/-.05/-.07/-.05/.00/.03/.07/.20/.40 |

CONSULTATION OF MEDICAL SPECIALIST

| | | | |
|---|---|---|---|
| A | VLVWKSP.K | = | DVWHAWS.KL |
| A | VLHHKSP.K | = | TBSSPWS.KL+TVWZKWS.KL |

PRESCRIPTIONS MEDICAL SPECIALIST

| | | | |
|---|---|---|---|
| A | VLVRSSP.K | = | (PATN_SP.K+PATN_ZK.K)*AVRSSPPC.K |
| N | AVRSSPPC | = | 2.2 |
| A | AVRSSPPC.K | = | 2.2+SPVRSDR.K+ZKVRSDR.K |
| A | SPVRSDR.K | = | TABHL(A5,DRUKESP.K,40,60,2.5) |
| T | A5 | = | -.03/-.05/-.07/-.05/-.03/.00/.03/.1/.2 |
| A | ZKVRSDR.K | = | TABHL(A6,BZGREZK.K,.70,.90,.025) |
| T | A6 | = | -.1/-.07/-.04/-.02/.0/.02/.04/.07/.1 |

CLINICAL TREATMENTS

| | | | |
|---|---|---|---|
| A | VLKVRZL.K | = | PATN_ZK.K*AKVRZKPC.K |
| N | AKVRZKPC | = | .24 |
| A | AKVRZKPC.K | = | .24+ZKKVRDR.K |
| A | ZKKVRDR.K | = | TABHL(A8,BZGREZK.K,.70,.90,.025) |
| T | A8 | = | .007/.004/.002/.001/.000/-.001/-.002/-.004/-.007 |

POLY-CLINICAL TREATMENTS

| | | | |
|---|---|---|---|
| A | VLPVRZL.K | = | PATN_SP.K*AKVRSPPC.K |
| N | APVRSPPC | = | .09 |
| A | APVRSPPC.K | = | .09+SPKVRDR.K |
| A | SPKVRDR.K | = | TABHL(A9,DRUKESP.K,40,60,2.5) |
| T | A9 | = | .012/.01/.007/.005/.003/.00/-.003/-.005/-.01 |

## COST OF MEDICAL PRODUCTION

| | | | |
|---|---|---|---|
| A | KOABOHA.K | = | PRABOHA*VLABOHA.K |
| A | KOCONHA.K | = | PRCONHA*VLCONHA.K |
| A | KOVRSHA.K | = | PRVRSHA*VLVRSHA.K |
| A | KOKVRZK.K | = | PRKVRZK*VLKVRZK.K |
| A | KOPVRSP.K | = | PRPVRSP*VLPVRSP.K |
| A | KOVRSSP.K | = | PRVRSSP*VLVRSSP.K |
| A | KOVWKSP.K | = | PRVWKSP*VLVWKSP.K |
| A | KOHHKSP.K | = | PRHHKSP*VLHHKSP.K |
| A | KOLIGZK.K | = | PRLIGZK*VLLIGZK.K |
| A | KOTOTAL.K | = | KOABO.K+KOCONHA.K+KOVRSHA.K+KOKVRZK.K+ KOPVRSP.K+KOVRSSP.K+KOVWKSP.K+KOHHKSP.K+ KOLIGZK.K |

## EXTERNAL DEVELOPMENTS

GROWTH OF POPULATION

| | | | |
|---|---|---|---|
| A | PATN_TO.K | = | PATN_TH.K+PATN_HA.K+PATN_WH.K+PATN_SP.K+ PATN_WS.K+PATN_WZ.K+PATN_ZK.K |
| R | GROEI.KL | = | .0001091*PATN_TO.K |

GROWTH OF CAPACITY

| | | | |
|---|---|---|---|
| C | AANT_HA | = | 102 |
| C | AANT_SP | = | 171 |
| C | AANT_BD | = | 1677 |

**PRICES (IN GUILDERS)**

| | | | |
|---|---|---|---|
| C | PRCONHA | = | 25.37 |
| C | PRABOHA | = | 1.923 |
| C | PRVRSHA | = | 10.70 |
| C | PRVWKSP | = | 67.70 |
| C | PRHHKSP | = | 27.00 |
| C | PRVRSSP | = | 18.75 |
| C | PRPVRSP | = | 82.50 |
| C | PRKVRZK | = | 177.50 |
| C | PRLIGZK | = | 321.00 |

# APPENDIX 8: CODEFORM A

| # of scene<br># of<br>sentence | cause | effect | sign | comments |
|---|---|---|---|---|
| | | | | |
| scene 1 | | | | |
| 1 | length-of-stay | number of patients in hospital | - | * in the short run<br>* no more than 5 % |
| 1 | number of patients in hospital | waiting list | - | |
| 2 | number of operation theatres | number of patients in hospital | + | |
| 2 | budget | number of patients in hospital | + | |
| 2 | number of nurses | number of patients in hospital | + | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

EXAMPLE

# APPENDIX 9: CODEFORM B

| # of scene<br># of<br>sentence | cause | effect | sign | comments |
|---|---|---|---|---|
| scene 1 | | | | |
| 1 | Measure (M) | 10 | | T, Q |
| 1 | 10 | 49 | - | |
| 2 | number of operation theatres (N) | 10 | + | |
| 2 | 24 | 10 | + | |
| 2 | number of nurses (N) | 10 | + | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

EXAMPLE

# APPENDIX 10: SUMMARY OF CODING RULES

*Establish causal relationships*

Since the recoding process is carried out per scene, read the entire scene first.
Use codeform A for the recoding process.

Assigning a sign to the relationship:
* if A increases and B increases (and vice versa), the sign is +
* if A increases and B decreases (and vice versa), the sign is -
* if A is affecting B, but it is not clear how, use O (=unclear)
* if it is stated that A is not affecting B, the relationship is denied (N)

Stay as closely as possible to the text. Interpret as little as possible.

Limit the number of implicit relationships. Implicit relationships are relationships one assumes to exist
to do right to the text. Code them by adding an 'i' behind the sign of the relationship.

If a concept has to be transformed for one reason or another, add (H) in the comment box for the concept.

Conditional relationships have to be coded as two separate relationships. The conditional part of the two
relationships has to be given a 'c' in the box in which the sign has to be written down as well.

Do not use terms as increase or decrease in the description of the concepts in the concept boxes.

Divide combined concepts (e.g. ratio's) into two or more concepts.

Either/or relationships should be coded as two separate relationships.

Relationships that are likely (e.g. A is likely to affect B), do not have to be coded differently from relationships in which no likelihood is expressed.

Repetitions should not be coded.

*Translating the concepts*

Read the scene once more.

Read the entire list (thesaurus) to make sure that you know all the standard concepts that can be selected from.

Determine as to whether the concept can be translated into the standard concept (measures should be coded as M).

If the concept used on the text can be translated into a standard concept but is somewhat more specific, an S should be written down in the relevant concept box behind the number of the standard concept. If the concept is more general than the standard concept, an A should be added.

If the concept used in the text seems to refer to more than one standard concept, and no decision can be made as to which of the two should be selected, write down a D and the concept used in the text. If you can make a decision regarding which of the two (or more) standard concepts should be selected, select the most appropriate standard concept, and write down an H behind it.

If no standard concept can be found,write down the concept used in the original text, and place an N in the concept box.

*Adding time-indicators and quantifications*

Read each sentence once more. If reference is being made to time (e.g. short term/long term), the relationship to which it applies should be coded as a (T) relationship. If a quantification is used (e.g. 10 per cent, 2000), then the concept to which it applies should be coded as a (Q) concept.

# APPENDIX 11: DICTIONARY

Concepts: from 000 to 252
Sections: from 1 to 8
1 = general practitioner section only
2 = medical specialist section only
3 = hospital section only
4 = all three sections
5 = medical specialist and hospital sections
6 = general practitioner and medical specialist sections
7 = general practitioner and hospital sections
8 = per definition a transition from one field to another
E := exogeneous, irrespective of the section
M := a measure, proposed by the participant

## *Concept number - section number - concept*

000 4 maatregel
001 3 aantal bedden
002 1 aantal huisartsen
003 2 aantal specialisten
004 4 aantal particuliere patienten
005 3 bezettingsgraad
006 1 drukte huisarts
007 5 drukte specialist
008 1 patienten op het spreekuur van de huisarts
009 2 patienten op het spreekuur van de specialist
010 3 patienten in het ziekenhuis
011 1 patienten die wachten op de huisarts
012 2 patienten die wachten op de specialist
013 3 patienten die wachten voor opname
014 1 consulteren huisarts
015 1 ontslagen worden door de huisarts
016 1 verwezen worden door de huisarts
017 1 terugbesteld worden door de huisarts
018 2 consulteren specialist
019 2 ontslagen door de specialist
020 2 (door)verwezen worden door de specialist
021 2 terugbesteld door de specialist
022 2 terugverwezen van specialist naar huisarts
023 3 opgenomen in het ziekenhuis
024 3 ontslagen uit het ziekenhuis
025 3 terugverwezen uit ziekenhuis naar specialist
026 3 terugverwezen uit ziekenhuis naar huisarts
027 1 wachttijd voor huisarts
028 2 wachttijd voor specialist
029 3 wachttijd voor opname
030 3 ligduur
031 1 aantal verrichtingen huisarts
032 2 aantal poliklinische verrichtingen specialist
033 2 aantal poliklinische verrichtingen specialist per contact
034 3 aantal klinische verrichtingen specialist
035 3 aantal klinische verrichtingen specialist per contact
036 5 aantal voorschriften specialist

037 1 ontslagpercentage huisarts
038 1 verwijspercentage huisarts
039 1 terugbestelpercentage huisarts
040 2 terugbestelpercentage specialist
041 2 specialist-huisarts terugverwijspercentage
042 3 ontslagpercentage ziekenhuis
043 1 abonnementstarief ziekenfondspatient
044 1E prijs van verrichtingen huisarts
045 5E prijs voorschrift specialist
046 3E prijs klinische verrichting
047 3 prijs ligdag
048 4 totale kosten
049 1 kosten abonnement huisarts
050 1 kosten voorschriften huisarts
051 1 kosten verrichtingen huisarts
052 2 kosten verwijskaarten
053 5 kosten voorschriften specialist
054 2 kosten poliklinische verrichtingen
055 3 kosten klinische verrichtingen
056 3 kosten verpleging ziekenhuis
057 1 inkomen huisarts
058 5 inkomen specialist
059 8 substitutie tweede-eerste lijn, vervangen dure zorg door goedkope zorg, verschuivingen verrichtingen specialist-huisarts
060 3 aantal ligdagen
061 3 operatiekamer capaciteit
062 1 huisartsendichtheid
063 2 specialistendichtheid
064 4 arbeidstijd
065 4 kwaliteit zorg
066 4 aandacht zieke, aandacht voor patient
067 4 bonus malus systeem
068 4 noodzaak van letten op kwaliteit
069 3 aantal afdelingen, aantal verpleegafdelingen
070 4 snelheid van doodgaan, levensduur
071 3 aantal verpleegkundigen, hoeveelheid verplegend personeel, hoeveelheid verpleegkundige hulp per bed
072 5 aantal topspecialisten
073 3 aantal specialisten die in priveklinieken gaan werken
074 3 aantal ziekenhuizen, nieuwbouw van ziekenhuizen
075 3 aantal verpleeghuizen
076 4 aantal regels van wat wel en wat niet mag
077 5 aantal waarnemingen
078 1 aantal maatschappelijk werkers
079 3 aantal diabetesverpleegkundigen
080 3 aantal stomadeskundigen
081 2 aantrekkelijkheid voor laagverdienende specialisten
082 2 aantrekkelijkheid voor veelverdienende specialisten
083 4 aantrekkelijkheid van de maatregel
084 4 aantrekken patienten
085 1 kosten thuishulp
086 4 accentverschuiving van kwalitatieve naar kwantitatieve aspecten
087 4 adherèntie, adherentie bevolking
088 4 administratieve onvolkomenheden
089 1 administratieve rompslomp huisarts, administratieve rompslomp, administratie huisarts

090 4 administratief/controlerende taak ziekenfonds, administratie ziekenfonds, uitbreiding van de verstrekkingenadministratie bij het ziekenfonds
091 4 afstemmen op maatregel
092 4 apparatuur
093 4 arbeidsvoorwaardenkwaliteit, rechten op grond van CAO
094 5 attitude specialist
095 4 beheersbaarheid van de kosten
096 5 bekendheid loon, bekendheid kosten loon specialisten landelijk
097 1 bekwaam voelen voor verrichting (huisarts), opbouw knowhow/vaardigheden huisarts
098 3 beleidsvrijheid ziekenhuis
099 4 beschadigende karakter van medische verrichtingen
100 3 beschikbare budget
101 2M bij het tarief budgetteren
102 5M bijstellen van het honoreringssysteem van de specialist
103 3 budget, budget ziekenhuizen
104 4 bureaucratie, bureaucratisering, overleg in bureaucratische sfeer
105 3 collegiale-begeleiding
106 4 centraliteit patient, patientgerichtheid van de zorg
107 1 collega's kruiswerk
108 1 complicaties thuis
109 3M compensatie in budgetsysteem
110 4 concurrentie, noodzaak van concurrentie
111 1 consumptiekosten huisarts, consumptiekosten
112 4 consulteren andere huisartsen, inter-collegiale consulten, onderling verwijzen
113 4 continuiteit
114 4 cursussen
115 3 dagverpleging, dagbehandeling
116 3 dekken van deze kosten (vaste kosten)
117 1 drempel verhogen voor abonnementstaken
118 4 dure diagnostiek
119 4 eerlijkheid van het systeem
120 1 eerstelijnsvoorzieningen, versterking eerste lijn, zorg eerste lijn
121 4 effecten gering opwaarts
122 3 efficiency ziekenhuisgebruik, efficiency bedgebruik
123 4 efficiency van werken, doelmatigheid, efficient met de tijd omgaan, efficiency
124 4 ellende patient
125 3 evenwichtiger allocatie van middelen binnen ziekenhuis
126 4 aantal vormen van zorg
127 5M externe toetsing
128 4 financiele prikkel (huisarts), prikkel, financiele bejag, stimulans, motivatie van specialisten, prikkel tot inkomenshandhaving, streven naar inkomenshandhaving, verlangen tot vergroting van het inkomen
129 5 financieel-economisch verantwoordelijk voor eigen toko, eigen baas, betrokkenheid specialist
130 3 flexibiliteit van reageren van management
131 1 fysiotherapie
132 1 gatekeepersfunctie huisarts
133 4 gedragsverandering, gedrag
134 4 gegarandeerdheid van het inkomen
135 1 geld kruiswerk
136 4 gevolgen van dien
137 4 gewenning
138 1 gezinssituatie
139 4 haastwerk, rust van werken
140 1 hoeveelheid werk kruiswerk
141 4 hoogleraren hebben kritiek indien niet aan alles gedacht wordt

142 5 in de hand houden van specialisten, afhankelijkheid van het management van specialisten, effectiviteit gezagstructuren, hoeveelheid gezag directie over handelen specialist, invloed van het ziekenhuis op specialist, beinvloeden handelen specialist
143 4 inspanningen, inspanning, moeite
144 1 instellingen voor thuiszorg
145 3 intensivering verpleegproces, relatieve zwaarte van de patienten, verzorging, zorg resterende verpleegdagen zwaarder, arbeidsintensiviteit per patient
146 3 intensive care
147 1 intensieve thuiszorg, thuiszorg, hoeveelheid thuiszorg, thuisverpleging, beroep op familie/mantelzorg,beroep op intensieve thuiszorg, druk thuiszorg, extra druk thuiszorg in brede zin, hoeveelheid hulp thuis nodig, familiehulp, hulp voor nazorg thuis, hoeveelheid hulp thuis nodig, aanvullende hulp thuis, vervangende hulp thuis
148 4 intensiteit praktijkvoering, intensiteit van behandelen/begeleiding van patient, arbeidsintensiviteit per patient, vasthouden van patienten
149 5M interne toetsing, intercollegiale toetsing, spiegelen
150 4 werkdruk
151 5 kosten niet-snijdende vakken
152 5 kosten snijdende vakken
153 3 inkomen ziekenhuis, baten
154 4 zorggerichtheid
155 4 zin om te bekwamen in die verrichtingen
156 4 werkzaamheden als invullen van medische kaarten, noodzakelijke activiteiten zoals bijvoorbeeld het invullen van medische kaarten
157 4E vraag, vraag naar medische zorg, latente zorgvraag, noodzakelijke hulpvraag
158 3 voorzieningen ziekenhuis
159 4 voorwaarden bij aanstellen
160 4 voorspelbaarheid kosten
161 4 voordeel
162 4 vinden van ziektes, ontdekken van onvolmaaktheden, vinden van kwalen
163 3 ruimte in het budget voor verpleegdagen
164 4E vergrijzing
165 3 vergoeding vaste lasten
166 3 vaste kosten, vaste lasten
167 3 variabele kosten in ziekenhuiszorg
168 4 tijdig diagnostiseren
169 4 tijd voor overige aanbod, tijd voor overige, tijd beschikbaar voor andere patienten, aandacht voor andere gebieden, andere taken huisarts, ruimte creëeren voor het doen van nieuwe taken
170 4 aantal effecten in eerste instantie
171 3 intramurale zorg
172 4M introductie van consult/jaarkaart
173 3 kapitaalverlies
174 3 kapitaalkosten
175 3 kapitaalslasten
176 4 kennis van de hoogte van de kosten, kennis van de kosten die hij veroorzaakt
177 4 koppeling niveau productie en niveau inkomen, inkomenseffect bij uitvoering van verrichtingen
178 5 kosten assistenten
179 4 kosten diensten, kosten waarneming
180 1 kosten eerste lijn, kosten eerste lijn door patienten thuis op te vangen
181 5 kosten tweede lijn
182 1 kosten huisartenhulp, huisarts, huisartsengeneeskunde
183 5 kosten specialistische hulp
184 5 kosten die de specialist als vervolgkosten genereert
185 3 kosten ziekenhuis/instelling, ziekenhuiskosten, kosten intramurale zorg, kosten binnen de instelling, exploitatiekosten

186 6 kosten extramurale zorg
187 1 kosten kruiswerk, kosten kruisvereniging
188 1 kosten gezinsverzorging, kosten gezinszorg
189 5 hulp specialist, hulp bieden
190 1 kosten fysiotherapie
191 4 kosten per mens
192 5 kosten per specialist
193 5 kosten per specialisme
194 4 kosten ziekenvervoer
195 4 kosten ziekenfonds
196 3 krapte budget ziekenhuis
197 4 kritischer indicatiestelling, indicatieverschuiving, scherpte van indicatie
198 3 kritischheid keuzemoment klinische/dagverpleging
199 4 kosten opleiding, opleidingskosten
200 4 kosten personeel
201 4 kosten materieel
202 4 kunnen bekwamen in die verrichtingen
203 4 liever eerder thuis
204 5E loonkosten specialist, loon specialist, prijs van een specialist in loondienst, salariskosten van de specialist in loondienst
205 4 marktmechanisme
206 4M maatregel aanbodzijde
207 4M maatregel vraagzijde
208 4E medische ontwikkeling, voorschrijdende techniek/ nieuwe behandelmethoden, mogelijkheden
209 4 medicalisering
210 3 mensen nodig, personeel tweede lijn, personeel ziekenhuis
211 4 mentaliteitsverandering
212 4 misbruik
213 3 mogelijkheden ziekenhuis
214 4 nascholing, tijd aan nascholing, eisen nascholing
215 4 niet voor elkaar onder doen, alles nakijken, -terughoudende geneeskunde, tijd om te onderzoeken
216 4 niet aanwezige diagnoses
217 1 omzet huisarts, omzet
218 3 ontslagdag
219 3 ontslagprocedure
220 5 open relaties tussen specialisten/specialismen
221 4 opleidingssituatie
222 4 opleiding aanpassen, eisen opleiding, aandacht binnen de opleiding voor verrichtingen, opleiding huisarts
223 4 tijd aan patienten besteden, tijd voor patient, tijd per patient, tijd voor nemen, langer mee bezig, tijd nodig voor onderzoek, tijd besteden voor deze verrichtingen, tijd voor het zelf doen van verrichtingen
224 3 opnamedag
225 3 opnames per bed
226 3 opnameprocedure
227 1 opvang
228 4 overheid moet kiezen
229 1 overige werkwijze huisarts
230 4E particuliere verzekering
231 4 patienten merken conflict
232 4E populatieveranderingen
233 4M preventie
234 4 productie, productiviteit, productiviteit per specialistproductie per specialist, aantal verrichtingen per specialist

235 4 productie/kwaliteitseisen
236 4 productieafspraken, maken van productieafspraken, afspraken over productie, afspraken maken over maximum productie per specialist
237 4 professionele status, professionele verantwoordelijkheid en autonomie
238 5M protocollering, afspraken over protocollen
239 3 restrictief beleid ziekenhuis
240 1 terugverwezen naar gezinsverzorging
241 3 opnamecapaciteit
242 8 samenwerking 1e en 2e lijn, goede samenwerking 1e/2e lijn, actief contact met ziekenhuis/specialist, afhankelijkheid/verbondenheid/loyaliteit met instelling, afspraken, afspraken maken met de huisarts over verrichtingen, afspraken tussen instelling en specialist omtrent lump-sum bedrag, afspraken over budget specialisten met ziekenfonds, betrokkenheid tussen specialist en ziekenhuis, conflicten met directie ziekenhuis, tegenstellingen specialist-overheid-ziekenhuis-verzekeraars, verantwoordelijkheid van specialist voor ziekenhuisexploitatie
243 4 snelheid van genezen
244 4 snelheid van klachten hebben
245 4E soort patienten
246 4 spanning, stress
247 3M specialistische hulp in ziekenhuisbudget, medeverantwoordelijkheid van specialisten voor ziekenhuisbudget
248 3 specialisten concentreren in 1 ziekenhuis
249 1 specialiseren van de huisartsen, verdiepen in verschillende deelgebieden
250 4 streven naar halen van norm
251 4 studietoelagen/onregelmatigheidstoeslagen
252 1 terugverwezen naar kruiswerk, naar kruiszorg

# APPENDIX 12: SPECIFYING EFFECTS OF THE BACKGROUND VARIABLES

The description of the specifying effects will start with an overview of the effects of the background variables and the amount of time invested in the program on individual changes in conceptualization. Following this, the effects of these potentially specifying variables on inter-individual changes will be discussed. Finally, the specifying effects on the outcome variable 'awareness of the others' point of view' will be presented.

## *Specifying effects on individual changes in conceptualization*

To examine the potentially specifying effects of '***age***', it was decided to divide the on-treatment group into a group of young (younger than 50 years of age) and group of older participants (50 years and older), to see as to whether differences could be found between these two groups with respect to the five research questions outlined before.

Regarding the individual changes in conceptualization, it was found that older people do better on almost all variables (both strategic and domain-specific) than younger people. Moreover, whereas young people decreased almost all their scores (that is posttest versus pretest), older participants managed to increase some of their scores. MANOVA analyses show that a significant difference between the two groups exists with respect to their individual domain-specific knowledge ($p=.08$).

To account for these differences, analyses were carried out to see as to whether any significant differences existed between young and old participants with respect to both background variables and variables concerning the evaluation of the program. These analyses show that no significant differences are found with respect to the evaluation variables, although older participants tend to appreciate the program a bit more (they score higher on the interesting, and useful aspects, and do not consider the program too long) than younger participants do. Moreover, older participants tend to appreciate the workbooks more than younger participants do, but dislike the small-groups activities more than relatively young participants do. As far as the other background variables are concerned, the only significant relationship that is found is the one (not surprisingly) between age and the number of years with the firm (Cramer's $V=.50$, $p<.05$). Note that older people do not do better than younger participants because of reasons such as higher education, or a higher score on investment of time, for no differences are found with respect to these variables.

The next variable that was found to have some specifying effect, is the background variable '***education***'. As shown in the figure below, people with a university degree outperform participants without such a degree on all variables with the exception of the number of time-indications. Not only do people with a university degree do much better (significantly on almost all strategic knowledge aspects (chaining: $p=.03$; connectivity: $p=.09$; feedbackloops: $p=.05$), and on the number of elements incorporated from the external model ($p=.03$) (ANOVA), they also improve on almost all variables (if their scores on pre- and posttest are compared to each other), with the exception of the endogeneity variable, where a small decrease is found. People without an university degree, however, have lower scores on the posttest than on the pretest on all variables.

To account for this striking difference, attention must be paid to variables such as amount of time invested in the program, age (for it already was found that older people do better than younger people do (in particular on the domain-specific dimension), years in the firm, age, and time-investment. No substantial relationships were found between level of education and any of the above-mentioned other background variables, although some association exists between educational background and department (Cramer's $V=.43$, $p=.10$).

Moreover, univariate analyses (ANOVA) show that people with an university degree differ from people without one, in that they dislike the workbook more than people without one do (p=.00), and appreciate the plenary sessions more than participants without a degree do. (p=.08). In addition to that, people without a degree look upon the program as relatively long (p=.09). However, they do not differ substantially from people with an university degree with respect to aspects such as useful, interesting, and easy.

| Age - group 1 (age < 50) versus group 2 (age >=50) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=12) | group 2 (N=6) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -8.67 (15.32) | -.83 (29.62) | 2 | .46 | .08* |
| exogeneity | -3.33 (10.09) | 8.33 ( 7.55) | 2 | .02* | |
| external model | -.25 ( 6.14) | 5.50 (13.82) | 2 | .23 | |
| feedbackloops | -.08 ( .79) | .33 ( .82) | 2 | .36 | .33 |
| connectivity | -.03 ( .13) | -.02 ( .22) | 2 | .97 | |
| chaining | .08 ( 2.15) | -.53 ( 1.43) | 1 | .54 | |
| time-indications | -.42 ( .79) | .17 ( .75) | 2 | .15 | |

| Education - group 1 (education=univ.ersity) versus group 2 (education=non-university) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=8) | group 2 (N=10) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -2.50 (26.69) | -8.90 (15.08) | 1 | .53 | .48 |
| exogeneity | 4.00 (13.84) | -2.20 ( 6.99) | 1 | .23 | |
| external model | 7.00 (10.97) | -2.60 ( 5.28) | 1 | .03* | |
| feedbackloops | .50 ( .54) | -.30 ( .95) | 1 | .05* | .24 |
| connectivity | .04 ( .46) | -.08 ( .15) | 1 | .09* | |
| chaining | .94 ( 2.34) | -.97 ( .96) | 1 | .03* | |
| time-indications | -.25 ( .46) | -.20 ( 1.03) | 2 | .40 | |

| Specialism - group 1 (one field) versus group 2 (more than one field) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=9) | group 2 (N=9) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -10.11 (23.36) | -2.00 (17.87) | 2 | .42 | .37 |
| exogeneity | -2.89 (12.05) | 4.00 ( 8.49) | 2 | .18 | |
| external model | 4.11 (10.94) | - .78 ( 7.40) | 1 | .28 | |
| feedbackloops | .00 ( 1.00) | .11 ( .78) | 2 | .79 | .84 |
| connectivity | -.07 ( .08) | .02 ( .21) | 2 | .23 | |
| chaining | -.35 ( .97) | .11 ( 2.60) | 2 | .62 | |
| time-indications | -.22 ( .67) | -.22 ( .97) | - | 1.00 | |

| Time-investment - group 1 (=< mean) versus group 2 (>mean) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=8) | group 2 (N=10) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -9.43 (17.01) | -3.91 (23.13) | 2 | .60 | .81 |
| exogeneity | -1.14 (11.19) | 1.64 (10.80) | 2 | .61 | |
| external model | 2.00 ( 5.16) | 1.46 (11.58) | 1 | .91 | |
| feedbackloops | .29 ( .76) | -.09 ( .82) | 1 | .39 | .50 |
| connectivity | .02 ( .12) | -.05 ( .18) | 1 | .40 | |
| chaining | .11 ( 1.11) | -.27 ( 2.34) | 1 | .70 | |
| time-indications | .14 ( .69) | -.46 ( .82) | 1 | .13 | |

| Years with the firm - group 1 (years =< 5) versus group 2 (years >5) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=6) | group 2 (N=12) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -12.33 (20.40) | -2.92 (20.85) | 2 | .38 | .07* |
| exogeneity | -7.33 (13.34) | 4.50 ( 6.75) | 2 | .02 | |
| external model | .33 ( 6.98) | 2.33 (10.63) | 2 | .68 | |
| feedbackloops | .50 ( .55) | -.17 ( .94) | 1 | .13 | .45 |
| connectivity | .00 ( .14) | -.04 ( .17) | 1 | .66 | |
| chaining | .87 ( 2.76) | -.62 ( 1.19) | 1 | .12 | |
| time-indications | .33 ( .52) | -0.17 ( .94) | 2 | .69 | |

| Policy experience - group 1 (60-100 % of time) versus group 2 (0-59 % of time) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=11) | group 2 (N=6) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -11.36 (16.34) | 2.29 (24.99) | 2 | .18 | .12 |
| exogeneity | 2.23 (11.29) | -2.86 ( 9.51) | 1 | .29 | |
| external model | - 1.18 ( 5.02) | 6.14 (13.04) | 2 | .11 | |
| feedbackloops | .00 ( 1.00) | .14 ( .69) | 2 | .75 | .99 |
| connectivity | -.03 ( .18) | -.02 ( .14) | 2 | .82 | |
| chaining | -.10 ( 2.40) | -.16 ( .92) | 1 | .95 | |
| time-indications | -.27 ( .91) | -.14 ( .69) | 2 | .75 | |

| Department - group 1 (financial-administrative) versus group 2 (medical) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=11) | group 2 (N=7) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -9.73 (15.61) | -.29 (27.08) | 2 | .36 | .01* |
| exogeneity | -5.09 ( 8.44) | 9.43 ( 7.48) | 2 | .00 | |
| external model | -.73 ( 6.19) | 5.43 (12.61) | 2 | .18 | |
| feedbackloops | -.18 ( .87) | .43 ( .79) | 2 | .15 | .37 |
| connectivity | -.05 ( .10) | .02 ( .23) | 2 | .40 | |
| chaining | -.47 ( 1.04) | .43 ( 2.85) | 2 | .35 | |
| time-indications | -.46 ( .82) | .14 ( .69) | 2 | .13 | |

| Gender - group 1 (male) versus group 2 (female) *on-treatment group* | | | | | |
|---|---|---|---|---|---|
| | group 1 (N=15) | group 2 (N=3) | | ANOVA | MANOVA |
| variables | posttest-pretest | posttest-pretest | best group | p-value | p-value |
| endogeneity | -1.87 (18.59) | -27.00 (19.98) | 1 | .05* | .03* |
| exogeneity | 2.93 ( 8.41) | -11.33 (15.04) | 1 | .03* | |
| external model | 1.87 (10.12) | .67 ( 5.69) | 1 | .85 | |
| feedbackloops | .07 ( .96) | .00 ( .00) | 1 | .91 | .99* |
| connectivity | -.03 ( .17) | -.04 ( .10) | 1 | .93 | |
| chaining | -.17 ( 2.09) | .11 ( .76) | 2 | .83 | |
| time-indications | -.27 ( .88) | .00 ( .00) | 2 | .62 | |

Regarding the background variable *'years with the organization'*, a somewhat more complicated pattern of differences is found. People who have been with the firm for more than 5 years, outperform people who have been at the firm for a shorter period of time on the domain-specific dimension (p=.07). However, with respect to the strategic dimension, people who have been with the firm for a shorter period of time, do much better than people who have been there for a longer period of time. As a matter of fact, the participants who just recently joined the firm, managed to increase their scores on three of the four strategic knowledge variables (no change was found on the connectivity variable).

The people who have been with the firm for more than five years considered the program to be easier than people who just joined the firm (p=.03). Moreover, they felt the program had been more useful (p=.03) and interesting (p=.04), than the younger participants did. The component of the program that they favoured most (compared to those who have worked for the firm for a shorter period of time) is the plenary session (p=.02).

As far as the association between the number of years with the firm and the other background variables is concerned, a positive association was found with age (Cramer's V=.50, p<.05) and department (Cramer's V=.89, p<.001) - the people who have been with the firm for more than five years come to a large extent from the medical services department, and are, on the whole, older.

Another interesting specifying effect was found when the background variable *'department'* was taken into account. As explained before, two departments took part in the program; a medical one and a financial-legal one. It was found that on all variables (domain-specific and strategic) the medical department did better than the financial-legal one. Moreover, the medical department succeeded in improving their scores on all variables (with the exception of the endogeneity variable, where no real difference was found). Note that only the domain-specific difference between the two departments can be considered as significant (p=.01).

To find out as to whether this difference between departments can be attributed to differences in background variables or differences in the evaluation of the program, analyses of variance were carried out.

The results of these analyses show that people from the medical department were of the opinion that the program had not been too long, in contrast to the people from the financial-legal department who felt it had been a somewhat long (p=.01). Moreover, it is interesting to see that the medical department appreciated the workbooks more than the financial-legal department did (p=.10), whereas the financial-legal department felt substantially more positive about the subgroup activities than the people from the medical department did (p=.06).

Regarding the other background variables, associations were found between department and age (Cramer's V=.88, p<.001), and years with the firm (Cramer's V=.89, p<.001). In other words, the people that work at the medical department, are older and have been with the firm for a longer period of time, which had already been shown in the description of the participants in Chapter Five.

Next, analyses were carried out on the basis of the background variable '*gender*'. As far as the domain-specific knowledge is concerned, males seem to outperform females (p=.03). However, one should take into account that only three females and 15 males did participate in the participative policy modelling sessions. Regarding the strategic knowledge dimension, the differences are somewhat mixed and less pronounced.

No clear relationships between gender and any of the components or aspects of the program were found which could shed some light on the differences between the male and female participants. Moreover no relationships between the other background variables and gender were found as well.

To assess as to whether any differences existed with respect to specialization, two categories had to be constructed first, one concerning the first echelon (physician/family doctor), and another concerning the second echelon, the medical specialist and hospital.

As shown in the table presented above, no significant differences between the two groups were found with respect to any of the outcome variables.

Finally, note that the people who spent less *'time'* on the program did do better on the strategic knowledge dimension than people who spent more time on it did. Although these differences are not significant, they are important for the people who spent less time on the program, not only do better than the people who spent more time on it, they also managed to *improve* on all four strategic knowledge variables. The time-investment variable was not found to be related significantly to any of the background variables, however, it was found to be significantly positively related to the aspect variables 'learned from it' ($r=.75$, $p<.01$) and 'useful' ($r=.57$, $p<.05$): those who spent more time on the program found the program to be more useful and claimed to have learned from it more than the people who spent less time on it.

Now that account has been given of the issue as to whether any (significant) difference exist between subgroups created on the basis of values for each of the background variables and the variable of time-investment, the question arises for those specifying variables where indeed interesting differences were found, as to whether they have been able to improve themselves significantly on any of the variables that make up the domain-specific and strategic dimensions. Because no significant results were found for the entire group of participants, it is only when significant differences are found between subgroups, that the possibility exists that significant improvements can be found for any of these subgroups.

The background variables that seem to qualify for further analyses in light of the analyses described above, hence are 'education' (in particular the group of participants with a university degree, for they improved on almost all variables), 'years with the firm' (a strong increase on the strategic knowledge variables was found for those with a limited number of years with the firm), and 'department' (the medical department did very well on most of the variables).

The analysis of variance analyses carried out to examine as to whether particular subgroups did indeed significantly improve their scores on either the domain-specific or strategic knowledge dimensions, however, show that in general no significant differences are found for the above mentioned three subgroups. The results of these analyses are presented below:

| | education=university (N=8) | | |
|---|---|---|---|
| variable | posttest-pretest | anova | manova |
| endogeneity | -2.50 (26.69) | .80 | .72 |
| exogeneity | 4.00 (13.84) | .44 | |
| external model | 7.00 (10.97) | .11 | |
| feedbackloops | .50 ( .54) | .04 * | .12 |
| connectivity | .04 ( .16) | .47 | |
| chaining | .94 ( 2.34) | .30 | |
| time-indications | -.25 ( .46) | .17 | |

| | department=medical (N=7) | | |
|---|---|---|---|
| variable | posttest-pretest | anova | manova |
| endogeneity | -.29 (27.08) | .98 | .05 * |
| exogeneity | 9.43 ( 7.48) | .02 * | |
| external model | 5.43 (12.61) | .30 | |
| feedbackloops | .43 ( .79) | .20 | .73 |
| connectivity | .02 ( .23) | .87 | |
| chaining | .43 ( 2.85) | .71 | |
| time-indications | .14 ( .69) | .60 | |

| | years with firm > 5 (N=6) | | |
|---|---|---|---|
| variable | posttest-pretest | anova | manova |
| endogeneity | -12.33 (20.40) | .20 | .46 |
| exogeneity | - 7.33 (13.34) | .23 | |
| external model | .33 ( 6.98) | .91 | |
| feedbackloops | .50 ( .55) | .08 * | .16 |
| connectivity | -.00 ( .14) | .97 | |
| chaining | .87 ( 2.76) | .47 | |
| time-indications | -.33 ( .52) | .18 | |

Summarizing the effects brought about by the specifying variables discussed in the present section on the *individual* changes in conceptualization, it has been shown that substantial differences exist between participants who have been with the firm for a long time, and those who just recently joined the firm, participants who have an university degree and participants who have not, and participants who work at the medical department and those who work at the financial-legal department. However, with respect to the subgroups that did significantly better than their counterpart (e.g. the medical department did do better than the financial-legal department), no significant improvements were found. Hence, one cannot conclude that an improvement of the domain-specific and strategic knowledge at any of the subgroups created on the basis these specifying variables was brought about by the participative policy modelling method.

### *Specifying effects on inter-individual changes in conceptualization*

With respect to the specifying effects of the background variables on the relationship between treatment (participative policy modelling) and degree of correspondence among the participants, tests of Bartlett were carried out to assess as to whether any significant differences in variance could be found between the groups constructed on the basis of the potentially specifying variables (e.g. groups of people with and without an university degree). The results of these analyses are the following (note that they were carried out on the on-treatment group - N=18):

| | education<br>university<br>non-university | policy<br>experience<br>much<br>little | department<br>medical<br>finan.-legal | specialization<br>1 field<br>more fields | gender<br>male<br>female |
|---|---|---|---|---|---|
| variables | p-value | p-value | p-value | p-value | p-value |
| endogeneity | .12 | .25 | .14 | .47 | .90 |
| exogeneity | .07 * | .66 | .75 | .34 | .25 |
| external model | .05 * | .01 * | .06 * | .29 | .39 |
| feedbackloops | .13 | .35 | .78 | .50 | - |
| connectivity | .81 | .47 | .03 * | .02 * | - |
| chaining | .02 * | .03 * | .01 * | .01 * | .38 |
| time-indications | .04 * | .49 | .65 | .31 | .17 |

- no values could be arrived at because of empty cells

| | time-investment<br>> mean<br>=< mean | years with the firm<br>more than 5<br>5 or less | workbook<br>> mean<br>=< mean | plenary session<br>>mean<br>=<mean | small-group<br>>mean<br>=<mean |
|---|---|---|---|---|---|
| variables | p-value | p-value | p-value | p-value | p-value |
| endogeneity | .43 | .96 | .38 | .46 | .85 |
| exogeneity | .93 | .07 * | .34 | .03 * | .22 |
| external model | .05 * | .32 | .21 | .14 | .01 * |
| feedbackloops | .57 | .21 | .62 | .46 | .02 * |
| connectivity | .24 | .64 | .86 | .98 | .20 |
| chaining | .07 * | .03 * | .11 | .00 * | .01 * |
| time-indications | .65 | .17 | .16 | .04 * | .24 |

| | useful<br>> mean<br>=< mean | interesting<br>more than 5<br>5 or less | easy<br>> mean<br>=< mean | length<br>>mean<br>=<mean | learned from<br>>mean<br>=<mean |
|---|---|---|---|---|---|
| variables | p-value | p-value | p-value | p-value | p-value |
| endogeneity | .36 | .53 | .30 | .99 | .18 |
| exogeneity | .46 | .40 | .86 | .23 | .29 |
| external model | .07 * | .11 | .09 * | .01 * | .09 * |
| feedbackloops | .99 | .85 | .63 | .14 | .09 * |
| connectivity | .49 | .65 | .10 | .07 * | .87 |
| chaining | .00 * | .04 * | .15 | .00 * | .04 * |
| time-indications | .81 | .46 | .31 | .05 * | .78 |

The potentially specifying variables (in this section the background variables are being dealt with only) where effects were found, are the variables 'department' and 'education'. The variance among the group of people with an university degree is different from the group of people with a non-university degree with respect to four of the total number of seven variables. As to the difference between the people from the medical department and the financial-legal department, note that the variances differ on the variables external model, connectivity, and chaining.

Having examined which of the potentially specifying variables do indeed specify the relationship between treatment and inter-individual change, the question as to whether significant improvements can be found for any of these subgroups, will be answered. Since no significant effects were found for the entire group, only those subgroups where significant effects were found in the analyses described in the previous section will be taken into account, for it is only with respect to those subgroups that significant differences potentially exist. As a consequence, we only need to examine subgroups that are based on the variables 'department' and 'education'. Moreover, as explained before, in order to qualify as an improvement in shared understanding, both the variance of the posttest has to be lower than the variance of the pretest, and the average score of the variable at hand on the posttest needs to be at least the same (preferably higher) as the score on the pretest. Hence, only the differences in variance between pre- and posttest of the group of people with an university degree, and the group of people who work for the medical department need to be examined, for those do show an increase rather than a decrease in their scores. The results of these analyses (using the test of Bartlett) have been presented in the following table:

| | education=university (N=8) | | | department=medical (N=7) | | |
|---|---|---|---|---|---|---|
| variables | pretest | posttest | p-value | pretest | posttest | p-value |
| endogeneity | 47.63 (23.77) | 45.13 (15.82) | .31 | 40.43 (15.29) | 40.14 (16.73) | .83 |
| exogeneity | 8.75 ( .79) | 12.75 (11.49) | .30 | 4.71 ( 4.96) | 14.14 (11.64) | .06 * |
| external model | 23.88 ( 9.73) | 30.88 (11.50) | .67 | 23.29 ( 7.74) | 28.71 (12.13) | .30 |
| feedbackloops | .38 ( .74) | .80 ( 1.13) | .30 | .43 ( .79) | .86 ( 1.07) | .48 |
| connectivity | 1.07 ( .18) | 1.11 ( .09) | .07 * | 1.04 ( .16) | 1.06 ( .13) | .61 |
| chaining | 7.01 ( 2.60) | 7.95 ( 2.22) | .70 | 6.39 ( 1.65) | 6.82 ( 2.20) | .50 |
| time-indications | .38 ( .74) | .13 ( .35) | .07 * | .29 ( .76) | .43 ( .54) | .42 |

Note than with respect to both specifying variables, shared understanding has been brought about at the connectivity variable for the group of people with an university degree only, because only there a significant reduction in variance and an increase in the score can be found. Since this may well be attributed to chance (particularly when alpha is set at .10, as is done in the present study), one should conclude that no shared understanding has been arrived at in both subgroups (the subgroup of people with an university degree and the subgroup of people from the medical department).

## *Specifying effects on awareness of the others' point of view*

Finally, the specifying effects of the background variables and time-investment on the fifth research question concerning the degree to which participants claim to have become aware of each others' point of view, needs to be examined. For this, analyses of variance (ANOVA) were carried out for each of the specifying variables. It turned out that neither the background variables, nor the time-investment variable did have any specifying effects on the degree to which the participants believed the participative policy modelling method had enabled them to acquire an awareness of each others' point of view. As a consequence, the very same conclusion drawn for the entire group holds for all the subgroups that can be created on the basis of the background variables and the variable of time-investment: participative policy modelling is able to bring about a moderate increase in the participants awareness of each others' point of view.

# APPENDIX 13: SPECIFYING EFFECTS OF THE EVALUATION OF THE PROGRAM

Having examined in detail the effects brought about by the potentially specifying variables 'time-investment' and 'background variables', account will be given of the influence the participant-based evaluation of the program may have on the relationship between the participative policy modelling methods and the dependent variables of the present study. The pattern that was followed in the description of the specifying effects of the background variables and time-investment will be used in the description of the effects of the participant-based evaluation as well. As a result, the description will start with the effects the potentially specifying variables may have on the individual changes brought about by the participative policy modelling method. Next, an overview will be given of the degree to which the inter-individual changes produced by the participative policy modelling method are affected by the participant-based evaluation. Finally, the specifying effects of these variables on the relationship between treatment and awareness of the other participants' point of view will be focused upon.

### *Specifying effects on individual changes in conceptualization*[174]

One of the most striking effects brought about by the specifying participant-based evaluation variables concerns the difference that is found between the group of people who considered the program to be somewhat short, and the people who found it relatively long. The people who felt the program had been somewhat short, improved their scores on all three domain-specific knowledge related variables, whereas the people who considered the program to be long, decreased their score on all three variables (in fact on all variables, both domain-specific and strategic). These differences were found to be highly significant. It almost seems as if the people who considered the program to be too long, had difficulty dealing with the complexity of the problem, as a result of which they may have become confused, which may have led to a decrease rather than an increase of their domain-specific knowledge scores.

Another interesting outcome concerns the differences that stem from the degree to which participants claim to have learned from the program. In contrast to what was expected, people who claim to have learned from it, score significantly lower on the strategic variables 'number of loops' and 'chaining'. Not only do they do worse than the group of participants who claim to have learned less, they also decrease their scores in an absolute sense, whereas the group of people who claim less to have learned from the program actually improve their scores. No significant differences between the people who claim to have learned more and those who claim to have learned less are found with respect to any of the domain-specific variables.

Finally, one should notice that substantial differences exist with respect to the domain-specific knowledge variables between people who say the program has been useful and people who have found the program less useful. The former group of people do improve their scores on all domain-specific variables substantially, whereas the latter group of people decrease their scores. It is only because of the relatively large standard deviations that none of these differences is significant.

[174]The analyses were carried out on the on-treatment group (N=18).

Differences between groups lower than and greater than the mean on evaluation of the program variables

| | length | | easy | | interesting | |
|---|---|---|---|---|---|---|
| | ANOVA | MANOVA | ANOVA | MANOVA | ANOVA | MANOVA |
| variables | p-value | p-value | p-value | p-value | p-value | p-value |
| endogeneity | .09 * | .08 * | .17 | .09 * | .97 | .35 |
| exogeneity | .07 * | | .25 | | .15 | |
| external model | .02 * | | .47 | | .31 | |
| feedbackloops | .79 | .67 | .06 * | .46 | .80 | .96 |
| connectivity | .52 | | .18 | | .53 | |
| chaining | .36 | | .30 | | .66 | |
| time-indications | .36 | | 1.0 | | .58 | |

| | useful | | learned from it | |
|---|---|---|---|---|
| | ANOVA | MANOVA | ANOVA | MANOVA |
| variables | p-value | p-value | p-value | p-value |
| endogeneity | .16 | .32 | .46 | .90 |
| exogeneity | .29 | | .93 | |
| external model | .09 * | | .60 | |
| feedbackloops | .70 | .78 | .06 * | .23 |
| connectivity | .83 | | .28 | |
| chaining | .53 | | .04 * | |
| time-indications | .26 | | .58 | |

| | workbook | | small-group session | | plenary session | |
|---|---|---|---|---|---|---|
| | ANOVA | MANOVA | ANOVA | MANOVA | ANOVA | MANOVA |
| variables | p-value | p-value | p-value | p-value | p-value | p-value |
| endogeneity | .04 * | .22 | .03 * | .06 * | .94 | .64 |
| exogeneity | .27 | | .16 | | .35 | |
| external model | .34 | | .08 * | | .93 | |
| feedbackloops | .27 | .30 | .46 | .59 | .43 | .79 |
| connectivity | .12 | | .49 | | .18 | |
| chaining | .09 * | | .80 | | .38 | |
| time-indications | .23 | | .15 | | .74 | |

### *Specifying effects on inter-individual changes*

Having examined the effect the participant-based evaluation of the program has had on the relationship between treatment and individual change in conceptualization, account must given to the effects of these variables on the inter-individual changes in conceptualization. For this, first of all, potentially differences between the variances of the subgroups created on the basis of the potentially specifying variables will be examined using the Bartlett test. On the basis of the outcomes of these analyses, comparisons between pre- and posttest variances will be made to assess as to whether a significant reduction in variance, indicating an increase in shared understanding, can be found for any of the subgroups were significant differences between the groups (rather than between pre- and posttest) were found.

As was shown in the table concerning the specifying effects on inter-individual changes, significant differences exist between the groups that are specified on the basis of the variables 'plenary session', 'small-group activities', 'length', and 'learned from it'. These figures mean that for instance the people who rated the plenary session below the mean have a significantly different value for the variance that exists among them with respect to their scores on variables such as exogeneity, chaining, and time-indications than people who rated the plenary session higher than the mean. In other words, the change in shared understanding differs significantly between both groups (the very same reasoning holds for the other three variables: small-group activities, learned from it, and length).

Next, the question as to whether for any of the subgroups created on the basis of the above-mentioned four specifying variables, a significant difference in the variance between pre- and posttest can be found, needs to be answered. To examine this, tests of Bartlett were carried out for all four specifying variables, however, only for those subgroups where no decrease in average score on the dependent variables was found, for in order to conclude that shared understanding has been arrived at, both the average scores must have maintained constant (or increase) and the variance should have been decreased. Consequently, it was decided to carry out these additional analyses for the following subgroups: the people with a score below the mean on the evaluation variables 'plenary session', 'length', and 'learned from it', and the people with a score higher than the mean on the variable 'small-group activities'. The results of these analyses are the following:

| | small-group < mean (N=9) | | | plenary session < mean (N=8) | | |
|---|---|---|---|---|---|---|
| variables | pretest | posttest | p-value | pretest | posttest | p-value |
| endogeneity | 46.50 (21.12) | 44.20 (14.53) | .28 | 52.00 (25.92) | 45.38 (12.18) | .06 * |
| exogeneity | 7.80 ( 6.99) | 10.90 (10.86) | .21 | 10.00 ( 9.15) | 8.00 (10.46) | .73 |
| external model | 24.40 ( 8.66) | 30.10 (10.27) | .62 | 27.13 (10.55) | 29.13 ( 7.42) | .37 |
| feedbackloops | .50 ( .71) | .70 ( 1.06) | .24 | .50 ( .76) | .75 ( .71) | .87 |
| connectivity | 1.06 ( .17) | 1.08 ( .10) | .19 | 1.05 ( .17) | 1.08 ( .12) | .34 |
| chaining | 6.86 ( 2.31) | 7.46 ( 2.33) | .92 | 7.38 ( 2.27) | 7.64 ( 2.30) | .97 |
| time-indications | .30 ( .68) | .20 ( .42) | .18 | .25 ( .46) | .00 ( .00) | - |
| | length < mean (N=9) | | | learned from it < mean (N=9) | | |
| variables | pretest | posttest | p-value | pretest | posttest | p-value |
| endogeneity | 36.78 (11.78) | 38.89 (16.29) | .38 | 44.56 (19.26) | 34.78 ( 9.63) | .07 * |
| exogeneity | 3.89 ( 4.29) | 9.11 ( 9.58) | .04 * | 5.67 ( 6.73) | 6.44 (10.43) | .24 |
| external model | 21.44 ( 7.83) | 28.33 (10.93) | .36 | 22.89 ( 8.79) | 23.33 ( 9.19) | .90 |
| feedbackloops | .78 ( .83) | .78 ( 1.09) | .46 | .22 ( .44) | .67 ( .71) | .20 |
| connectivity | 1.04 ( .17) | 1.03 ( .14) | .64 | 1.02 ( .14) | 1.04 ( .12) | .63 |
| chaining | 6.86 ( 2.41) | 7.09 ( 2.66) | .79 | 6.14 ( 2.38) | 6.94 ( 2.36) | .98 |
| time-indications | .11 ( .33) | .11 ( .30) | 1.00 | .64 ( 1.12) | .33 ( .71) | .22 |

In line with the evaluation of the effects brought about by the background variables and time-investment, one should conclude that no shared understanding has been arrived at in the subgroups 'people who found the program to be relatively short', 'people who claimed less they had learned from the program', 'people who appreciated the small-group activities very much', and 'people who did rate the plenary session below the mean'.

## *Specifying effects on awareness of the others' point of view*

Finally, the specifying effects of the participant-based evaluation of the program of the fifth research question must be examined. To do so, analysis of variance (ANOVA) have been carried out for each of the potentially specifying variables, to determine as to whether any significant differences could be found between the subgroups composed on the basis of their values for each of the potentially specifying variables. The results of these analyses are the following:

| | awareness |
|---|---|
| variables | anova |
| time-investment (<>mean) | .73 |
| specialism (1, more fields) | .61 |
| education (univ., non-univ) | .94 |
| years with the firm (=<5, >5) | .35 |
| policy experience (>59%,=<59%) | .33 |
| department (medical,finan-adm.) | .14 |
| gender (male, female) | .15 |
| age (<>mean) | .73 |

| | awareness |
|---|---|
| workbook (<>mean) | .57 |
| plenary sesion (<>mean) | .13 |
| small-group session (<>mean) | .05 * |
| useful (<>mean) | .14 |
| interesting (<>mean) | .01 * |
| easy (<>mean) | .15 |
| length (<>mean) | .19 |
| learned from (<>mean) | .57 |

The only significant differences that are found, are the ones between the groups that are created on the basis of the variables 'small group activities' (p=.05) and 'interesting' (p=.01). It means that people who claim that the small group activities have been interesting (they score higher than the mean), have a significantly lower score on the awareness variable (3.09 on average), than the people who scored lower than the mean on the interesting variable.

As a consequence, one should conclude that the participant-based evaluation of the program in terms of components and aspects does not specify significantly the relationship between treatment and awareness of each others' point of view.

# APPENDIX 14: WORST CASES VS BEST CASES

Another way to specify the relationship between treatment and effects, is to analyze as to whether any differences exist between people who did extremely well and people who did extremely bad with respect to the differences between pre- and posttest. In other words, the people who improved their score dramatically, and the people who were found to have reduced their scores substantially will be compared to see as to whether any interesting differences can be found between the two groups with respect to their background variables, the way in which they evaluated the program, and the amount time they spent on it. Differences between these two groups may give us additional information on what it takes to either improve or deteriorate one's conceptualization by taking part in the program. Because the best cases/worst cases analysis is based on only a relatively few number of cases, a description of the outcomes will be given rather than statistically testing for difference between the two groups. Moreover, since the number of people in the worst cases and best cases groups is very small, no account will be given of the inter-individual changes brought about at these groups, no attempt will be made to assess as to whether a difference exists in the shared understanding arrived at in these groups. However, differences with respect to the degree to which the worst and best cases groups claim to have become aware of each others' point of view as a result of their participation will be discussed in the present section.

The first comparison concerns the differences that exist between participants who increased and the participants who decreased their scores on the domain-specific variables 'endogeneity' and 'exogeneity' (only those who increase their scores on both of the two variables qualify for the best cases groups and those who decrease on both variables, qualify for the worst cases group). Since we would also like to know as to whether any differences exist between those two groups with respect to the strategic knowledge variables (do people with low improvement also do bad on the strategic knowledge dimension?) scores on these variables will be compared as well.

The second comparison will be made between people who did well and people who did bad on the strategic knowledge dimension, for it may well be that different factors play a role in this distinction than the factors that are important to the distinction between best and worst domain-specific knowledge cases. To distinguish between best and worst cases, participants either had to increase their scores on connectivity and chaining (to qualify for the best cases group), or decrease their scores on both variables (to qualify for the worst cases group). In the comparisons between the worst and best strategic knowledge groups, domain-specific variables have been included as well. The results of these comparisons with respect to the dependent variables (individual change) are depicted below:

| Worst/best cases comparisons | | | | |
|---|---|---|---|---|
| | worst cases group (N=3) | | best cases group (N=3) | |
| *domain-specific* | domain-specific knowledge | strategic knowledge | domain-specific knowledge | strategic knowledge |
| endogeneity | -20.25 | -18.00 | 23.67 | 1.20 |
| exogeneity | -11.25 | -1.00 | 9.67 | 8.42 |
| external model | -2.75 | -.17 | 15.67 | 8.80 |
| *strategic* | | | | |
| feedbackloops | .00 | -.50 | .67 | .62 |
| connectivity | -.02 | -.18 | .16 | .12 |
| chaining | -.02 | -1.09 | 1.59 | 1.96 |
| time-indications | -.75 | -.17 | .00 | -.20 |

Regarding the differences between the groups with respect to the domain-specific and strategic knowledge, the people who score low on the domain-specific dimension will be compared to the people who are bad on the strategic dimension, and vice versa to see as to whether different qualities are required to do well (or bad) on the individual dimensions.

It is shown in the above figure, that those who do bad on the domain-specific dimension also do bad on the strategic dimension, in that they decrease their scores on all variables except for the number of feedbackloops where no change is found. This is not to say that they are the same people as the people who fall in the 'worst strategic knowledge cases'. The people who did bad on the strategic dimension (that is, those whose score went down), also did bad on the domain-specific dimension -those who decrease their strategic knowledge score also decrease their domain-specific knowledge.

If we compare both worst cases groups, we see that, as expected, the domain-specific worst cases do worse than the strategic knowledge worst cases group on the domain-specific variables, and that vice versa, the strategic worst cases group is doing worse on the strategic variables than the domain-specific worst cases group is doing.

However, if we look at the best cases groups, we see that although both best cases groups (the domain-specific best and the strategic knowledge best) increase their scores on all variables (with the exception of the time-indications variable), the people who do best on the domain-specific dimension outperform the strategic best cases group on almost all variables, that is, on both domain-specific and strategic dimensions[175]. In other words, the people who improve on the domain-specific dimension also improve on the strategic dimension, whereas the people who improve most on the strategic knowledge dimension improve substantially less on the domain-specific knowledge dimension than the people who improve most on domain-specific knowledge do. One should bear in mind that this does not mean that people who have a high score on the domain-strategic dimension, also have a high score on the strategic dimension, but that those who improved substantially on the first, also improved on the second one, and vice versa.

[175]On first thought, this would seem impossible, for if this would be the case, the best cases of the domain-specific best cases group would have to be included in the group of best strategic knowledge cases (indeed 2 of the 5 subjects of both best cases groups are identical). However, close examination of the data show that on the chaining variable, the strategic group outperforms the best domain-specific cases group.

The differences that exist between these groups with respect to the evaluation of the program, background variables, and the amount of time invested by the participants can be displayed as follows:

| Worst/best cases comparisons | | | | |
|---|---|---|---|---|
| | domain-specific | | strategic | |
| *aspects* | worst (N=4) | best (N=3) | worst (N=6) | best (N=5) |
| length | 3.67 | 2.38 | 3.01 ◄ | ► 2.68 |
| easy | 3.52 | 3.33 | 3.15 | 3.39 |
| interesting | 3.58 | 3.67 | 4.06 ◄ | ► 3.58 |
| useful | 3.28 ◄ | ► 3.75 | 3.57 | 3.58 |
| learned from it | 3.10 | 3.05 | 3.24 ◄ | ► 2.85 |
| *components* | | | | |
| small-group session | 4.10 ◄ | ► 3.57 | 3.69 | 3.64 |
| plenary session | 3.54 | 3.34 | 3.46 | 3.39 |
| workbook | 3.11 ◄ | ► 3.42 | 3.20 | 3.34 |
| *background* | | | | |
| policy experience | 2.75 | 2.67 | 3.50 | 3.40 |
| years with the firm | 2.25 ◄ | ► 3.67 | 3.50 | 3.40 |
| age | 1.50 ◄ | ► 3.33 | 2.30 ◄ | ► 2.80 |
| total time invested | 13.1 | 15.8 | 15.0 | 13.20 |
| awareness | 3.39 | 3.19 | 3.33 | 3.11 |

◄─► Difference between the two groups of cases >= .30

Notice that only background variables of ordinal or interval level have been included in the above figure for simplification purposes. Nominal variables such as department, education, and gender will be included in the description of the four groups, whenever interesting differences are found on those variables.

The group of participants who improve less on the domain-specific dimension (as a matter of fact, they decrease their scores), can be characterized as a group consisting of people from the financial-administrative department, both males and females, of mixed educational background, who have been with the firm for a considerably shorter period of time than people who do well on the domain-specific knowledge dimension have. The people who did do well on the domain-specific dimension, are all males, with an university degree, and working for the medical services department. They are on average much older than the people from the worst cases group.

Regarding the differences in appreciation of the program, it is noticeable that those who did worst on domain-specific knowledge, did appreciate the small-group sessions much more than those who did do best, but scored lower on the workbook variable. Moreover, the people who did do best, found the program to be more interesting than those who did do worst. Finally, note that with respect to the perceived length of the program, the group of people who scored worst, felt the program had been somewhat long (this is not due to the fact that they found the program to be difficult or uninteresting, for their scores on those variables are not considerably different from the best cases group).

If we compare the people who did worst and best on the strategic knowledge dimension, we can see that those who did worst on the strategic dimension come from both departments, are both male and female, and have a mixed educational background. Moreover, they are considerably younger than those who do well on the strategic dimension. The people who did well on the strategic knowledge dimension come from both departments, and have all (except for one participant) an university background.

Regarding the way in which the groups of best and worst cases looked upon the program and the amount of time spent on it, it is surprising to see that those who did best rated the program as less interesting as those who did do worst (may be because they felt they had not learned much from it for the people who did do best, felt the program had not provided them with a lot of knowledge). The people who did do best on the strategic knowledge dimension, also found the program to be a little bit short, in contrast to those who did do worst, who felt that the length of the program was about right. Note that no correlations exist between the participant-based evaluation variable 'length' and the variables 'learned from it' and 'easiness'. People who have found the program a bit long are not the people who felt the program was either too difficult or too easy, or the people who believed they had not learned from it at all. Moreover, no relationship was found to exist between the amount of time invested in the program, and the scores on the aspect variable 'length': the people who spent more time on the program, not necessarily are the people who felt the program had been too long or vice versa.

With respect to the average score on awareness of each others' point of view, note that the people belonging to the best cases groups (both the group of people who are best on the domain-specific and strategic dimensions) score considerably lower on this variable than the people who belong to the worst cases groups do. As such, it follows the findings discussed above with respect to the relationship 'learned from it according to the cognitive map indicators' and 'learned from it according to the questionnaire items': in both cases those who do well on the cognitive map criteria, claim to have learned little, with respect to the problem at hand *and* the people who also participated to the program.

A possible explanation for the fact that those who do well on the domain-specific and strategic knowledge dimensions, claim to have learned little about the other participants (i.e. they have a relatively low score on the awareness variable), may be that people either concentrate on the knowledge required to deal with the problem, or focuses on the 'social' interaction that is taking place during the sessions by means of which knowledge of the other participants can be acquired. The reason for this may be that there simply is not enough time to do both or that people who do not enjoy, say the 'knowledge of the problem'-aspect of the program, concentrate on the other aspect: the social interaction that is taking place during the sessions. Some evidence for the last reason may be that people who did worst (this goes for both groups), enjoy the small-group activities (where intensive social interaction is taking place) more than people who did do best. In addition to this, the people who improved most, enjoyed the workbooks (in which a lot of problem-related knowledge is given to the participants) more than people who did not improve much. Irrespective of the reasons why, it is interesting to see that those who do well on the individual conceptualization of the problem, are not the ones who claim to have become aware of the other participants' point of view.

Finally, note that those who did best on the domain-specific dimension, put in substantially more hours than those who did worst on that very same dimension. The inverse relationship holds for the strategic knowledge group: those who did best are the people who spent less time on the program. It almost seems as if putting in more time will help you increase your domain-specific knowledge, but helps you decrease your strategic knowledge. However, since no clear negative correlation is found between domain-specific knowledge scores and strategic knowledge scores, one cannot conclude that increasing one's domain-specific knowledge score will result in a decrease in one's strategic knowledge score, or the other way around.

# CURRICULUM VITAE

Luc Verburgh (1961) completed his secondary education in1980 (Atheneum B) at the Maurick College in Vught, the Netherlands. He studied Philosophy in Nijmegen, the Netherlands and Applied Organizational Psychology at the University of Wales Institute of Science and Technology in Cardiff (UK). From 1988 until 1992, he worked as a research assistant at the Department of Methodology in Nijmegen. Besides working on the evaluation of the effects of participative policy modelling, he was involved in the development of a management game for building societies. In 1990 he started working as a freelance consultant for the European Professional Education division of Arthur Andersen & Co. S.C. As part of that job, he provided Management Development training in the Netherlands and abroad. In 1992, Luc Verburgh joined Andersen Consulting in the Hague, where he is working as a consultant in the Change Management practice of the Financial Services Industry group.

Many organizations are confronted with ill-structuredness and complexity. To support organizations in dealing with this complexity, use can be made of a participative version of system dynamics modelling. This doctoral thesis provides a detailed description of how to evaluate the effects of participative policy modelling. It assesses the degree to which people from a health care insurance organization change the way in which they conceptualize the complex problem of continuously rising health care costs. In order to interpret the results of this evaluation, a new theoretical framework is introduced. Central in this framework is the notion that to understand the complexity of the problem one is confronted with, a change in perspective is needed. It is only by changing one's perspective from an outsider's (observer's) to an insider's (participant's) point of view and vice versa, that a thorough understanding of the structure and dynamics of the problem can be acquired and that structurally sound solutions to the problem can come to mind.

The thesis may be of interest to those who would like to know more about participative policy modelling and its applicability. Moreover, it provides an introduction to the structure of the Dutch Health Care System, as well as an overview of the system's dynamic behaviour. In addition, this thesis can be used as a reference for those who would like to know more about evaluating changes in the way in which people view a particular policy problem.

The thesis was written in partial fulfillment of the requirements for the Degree of Doctor in the Social Sciences. At the time the research was carried out, Luc Verburgh held a position as research assistant at the Department of Methodology at the University of Nijmegen, the Netherlands. He is currently a consultant at Andersen Consulting in the Hague, working at the Change Management practice of the Financial Services Industry group.

ISBN 90 9006953 4